

Decoupling on the Wiener Space, Related Besov Spaces, and Applications to BSDEs

Stefan Geiss

Juha Ylinen

Author address:

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF JYVÄSKYLÄ,
P.O. BOX 35, FINLAND

E-mail address: stefan.geiss@jyu.fi

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF JYVÄSKYLÄ,
P.O. BOX 35, FINLAND

E-mail address: juha.m.ylinen@jyu.fi

Contents

Chapter 1. Introduction	1
1.1. Background	1
1.2. Outline of the main ideas	1
1.3. Notation	7
Chapter 2. A General Factorization	9
2.1. The operators \mathcal{C}_S	9
2.2. The operators \mathcal{C}_S^M as extension of \mathcal{C}_S	14
Chapter 3. Transference of SDEs	17
3.1. Setting	17
3.2. Results	17
Chapter 4. Anisotropic Besov Spaces on the Wiener Space	21
4.1. Classical Besov spaces on the Wiener space	21
4.2. Setting	23
4.3. Definition of anisotropic Besov spaces	24
4.4. Connections to real interpolation	27
4.5. The space $\mathbb{B}_p^{\Phi^2}$	29
4.6. An embedding theorem for functionals of bounded variation	35
4.7. Examples	36
Chapter 5. BMO-Martingales	41
5.1. BMO-martingales	41
5.2. Reverse Hölder inequalities	44
5.3. Fefferman's inequality and $\text{BMO}(S_{2\theta})$ spaces	46
Chapter 6. Applications to BSDEs	53
6.1. The setting	53
6.2. Stability of BSDEs with respect to perturbations of the Gaussian structure	55
6.3. On classes of quadratic and sub-quadratic BSDEs	60
6.4. Settings for the stability theorem	67
6.5. On the L_p -variation of BSDEs	73
6.6. An a-priori estimate for BSDEs	79
6.7. Concluding remarks	83
Appendix A. Technical Facts	85
Bibliography	95

Abstract

We introduce a decoupling method on the Wiener space to define a wide class of anisotropic Besov spaces. This class of Besov spaces contains the traditional isotropic Besov spaces obtained by the real interpolation method, but also new spaces that are designed to investigate the L_p -variation of backwards stochastic differential equations where the generator might be of quadratic type. Our decoupling method is based on a general distributional approach not restricted to the Wiener space.

2010 *Mathematics Subject Classification.* Primary 60H07, 60H10, 46E35.

Key words and phrases. Anisotropic Besov spaces, Decoupling on the Wiener Space, Backward Stochastic Differential Equations, Interpolation.

The authors were supported by the project "Stochastic and Harmonic Analysis, interactions, and applications", No. 133914, of the Academy of Finland. The second author was partly supported by the Vilho, Yrjö and Kalle Väisälä foundation of the Finnish Academy of Science and Letters.

CHAPTER 1

Introduction

1.1. Background

A backward stochastic differential equation (BSDE) is an equation of type

$$(1) \quad Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s,$$

where $T > 0$ is a fixed time horizon, $W = (W_t)_{t \in [0, T]}$ is a d -dimensional Brownian motion, $\xi : \Omega \rightarrow \mathbb{R}$ a given \mathcal{F}_T -measurable *terminal condition*, and

$$f : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$$

a given predictable random generator which might be non-Markovian. Given the data (ξ, f) , one looks for adapted solution processes (Y, Z) . Backward stochastic differential equations have a wide range of applications, for example in stochastic control and, more generally, in stochastic modeling. In the case of a Markovian generator, where the randomness comes from a forward diffusion, there is an important and extremely useful connection to non-linear partial differential equations of parabolic type, the so-called (non-linear) Feynman-Kac theory. Two seminal papers in this theory were the work of Bismut [11], and Pardoux and Peng [62].

The simulation of BSDEs is an important topic and subject to active research. To setup simulation schemes one needs an approximation theory for BSDEs, for example to find optimal time-grids or to obtain upper and lower rates for the speed of convergence of these schemes measured in an appropriate way. To investigate these approximation properties it is more or less mandatory to understand the variational properties of the solution (Y, Z) , i.e. the behavior of

$$(2) \quad \|Y_t - Y_s\|_p \quad \text{and (say)} \quad \left\| \left(\int_s^t |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p$$

for all $0 \leq s < t \leq T$ and an appropriate range of $p \in (0, \infty)$.

1.2. Outline of the main ideas

In these notes we develop an approach to estimate the variations from (2) in terms of the regularity of the data (ξ, f) , where the regularity is a fractional smoothness expressed in terms of Besov spaces. Our approach is based on an anisotropic decoupling of the Wiener space. Recently this decoupling was already successfully used in [35, 36] and constitutes one of the few approaches to estimate variational properties of non-Markovian backwards equations using only knowledge of the initial data. Let us explain the basic line of ideas to motivate the structure of these notes.

If the generator in our BSDE vanishes, i.e. $f \equiv 0$, then one has that

$$Y_t = \mathbb{E}(\xi | \mathcal{F}_t).$$

Therefore, in the case $f \not\equiv 0$ the map

$$G_t^f : \xi \rightarrow Y_t$$

can be interpreted as some kind of generalized *non-linear conditional expectation* along the generator f (see [63, 26, 64] for the notion of g -expectation and nonlinear expectations). Hence, our notion of regularity should be stable with respect to this non-linear map G_t^f . Moreover, since

$$\|Y_t - Y_s\|_p \leq \|Y_t - \mathbb{E}(Y_t | \mathcal{F}_s)\|_p + \|\mathbb{E}(Y_t | \mathcal{F}_s) - Y_s\|_p$$

and since $\|\mathbb{E}(Y_t | \mathcal{F}_s) - Y_s\|_p$ can be handled by 'standard' methods, the main question consists in investigating the behavior of $\|Y_t - \mathbb{E}(Y_t | \mathcal{F}_s)\|_p$ for $s \uparrow t$. It turns out that this behaviour corresponds to a notion of fractional smoothness in L_p of the random variable Y_t . The crucial point here is that

$$(3) \quad \|Y_t - \mathbb{E}(Y_t | \mathcal{F}_s)\|_p \sim \|Y_t - Y_t^{(s,t)}\|_p$$

for $p \in [1, \infty)$, where $Y_t^{(s,t)}$ is a decoupled version of Y_t in the sense explained below. Therefore we proceed as follows:

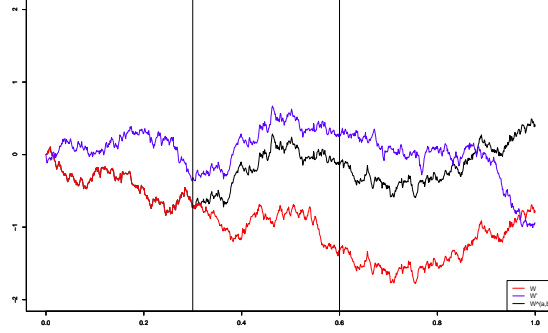
- (a) In Chapter 2 we introduce a factorization on which our decoupling method is based on,
- (b) in Chapter 3 we apply this factorization method to stochastic differential equations,
- (c) in Chapter 4 the decoupling and the corresponding Besov spaces on the Wiener space are introduced and investigated,
- (d) in Chapter 5 we provide some tools about BMO spaces and reverse Hölder inequalities important for the treatment of non-Lipschitz BSDEs,
- (e) and in Chapter 6 we finally apply the results of Chapters 2, 3, 4, and 5 to BSDEs.

We proceed with some exemplary ideas and results obtained in this article:

Chapters 2 - 4: The decoupling to obtain $F^{(a,b]}$ from a random variable $F : \Omega \rightarrow \mathbb{R}$ on the Wiener space is done as follows: We start with a Wiener space built on a d -dimensional Brownian motion $W = (W_t)_{t \in [0, T]}$. Then we take a copy of this Wiener space, denote the corresponding Brownian motion by $W' = (W'_t)_{t \in [0, T]}$, and form the canonical product space carrying the $2d$ -dimensional Brownian motion $((W_t, W'_t))_{t \in [0, T]}$. But the pair (W, W') of Brownian motions is not the one we are interested in in the sequel. Instead, we take (for example) an interval $(a, b] \subset (0, T]$ and consider the mixed Brownian motion $W^{(a,b]} = (W_t^{(a,b]})_{t \in [0, T]}$ where the increments on the interval $(a, b]$ from W are replaced by the increments of the independent copy $W' = (W'_t)_{t \in [0, T]}$, i.e. we define

$$W_t^{(a,b]} = \begin{cases} W_t & : 0 \leq t \leq a \\ W_a + W'_t - W'_a & : a \leq t \leq b \\ W_a + (W'_b - W'_a) + (W_t - W_b) & : b \leq t \leq T \end{cases}.$$

In other words, the Gaussian structure on $(a, b]$ is replaced by an independent copy:



Now the random variables F from the original Wiener space built on W are extended to the product space carrying (W, W') and are transformed by a functional mapping $F \rightarrow F^{(a,b]}$ along the same map as $W \rightarrow W^{(a,b]}$ is transformed.

After we have introduced the decoupling method, our next step consists in observing that one can define anisotropic Besov spaces by imposing Hölder type conditions on a random variable $\xi \in L_p$ like

$$\|\xi - \xi^{(a,b]}\|_p \leq c \alpha(a, b)$$

for all $0 \leq a < b \leq T$ and an appropriate weight function $\alpha(\cdot, \cdot)$. These anisotropic Besov spaces are part of a wider class of spaces containing the traditional Besov spaces obtained by the real interpolation method.

To explain a prototype of these Besov spaces, let us assume $p, r \in [2, \infty)$ and $\xi \in L_p$. In Chapter 4 we use

$$\alpha_r(a, b) := \sqrt[r]{b-a}$$

to define $\xi \in \mathbb{B}_p^{\Phi_r}$ provided that

$$\|\xi\|_{\mathbb{B}_p^{\Phi_r}}^p := \mathbb{E}|\xi|^p + \|\xi\|_{\Phi_{r,p}}^p < \infty \quad \text{with} \quad \|\xi\|_{\Phi_{r,p}}^p := \sup_{0 \leq a < b \leq T} \left| \frac{\|\xi - \xi^{(a,b]}\|_p}{\sqrt[r]{b-a}} \right|^p.$$

The case $r = 2$ is treated by Theorem 4.19 and includes the following situation, where $\mathbb{D}_{1,2}$ stands for the Malliavin Sobolev space and $D\xi$ for the Malliavin derivative:

THEOREM 1.1. *One has $\mathbb{B}_2^{\Phi_2} \subseteq \mathbb{D}_{1,2}$. Moreover, for $p \in [2, \infty)$ and $\xi \in \mathbb{D}_{1,2} \cap L_p$ it holds*

$$\|\xi\|_{\Phi_{2,p}} \sim_c \sup_{0 \leq a < b \leq T} \left\| \left(\frac{1}{b-a} \int_a^b |D_s \xi|^2 ds \right)^{\frac{1}{2}} \right\|_p,$$

where $c > 0$ depends on p only. In particular, for $p = 2$ we have that

$$\|\xi\|_{\Phi_{2,2}} \sim_c \text{esssup}_{s \in [0, T]} \|D_s \xi\|_2.$$

The impact of Theorem 1.1 (Theorem 4.19) is at least twofold: Firstly, we can access the Malliavin derivative by the spaces $\mathbb{B}_p^{\Phi_2}$ without using the derivative explicitly.

Secondly, the above theorem can be localized by replacing $\|\xi\|_{\mathbb{B}_p^{\Phi_2}}^p$ with

$$\mathbb{E}|\xi|^p + \sup_{A \leq a < b \leq B} \left| \frac{\|\xi - \xi^{(a,b)}\|_p}{\sqrt{b-a}} \right|^p$$

for some $0 \leq A < B \leq T$. Here ξ does not need to belong to $\mathbb{D}_{1,2}$ anymore.

The case $r = 4$ turns out to be relevant for the local time of a Brownian motion, for example represented by

$$L_t^\alpha = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t \chi_{(\alpha-\varepsilon, \alpha+\varepsilon)}(W_s) ds \text{ a.s.}$$

Here we prove in Corollary 4.27 that for all $\alpha \in \mathbb{R}$ and $p \in (1, \infty)$ one has that

$$L_T^\alpha \in \mathbb{B}_p^{\Phi_4} \setminus \left[\bigcup_{r \in [2,4)} \mathbb{B}_p^{\Phi_r} \right].$$

Chapter 5: Given a continuous BMO-martingale M and its Doléan-Dade exponential $\mathcal{E}(M)$, we introduce the sliceable numbers $\text{sl}_N(M)$ in Definition 5.2 that measure the distance of M to \mathbb{H}_∞ . Denoting by $\mathcal{RH}_\beta(\mathcal{E}(M))$ the constant in the reverse Hölder inequality for $\mathcal{E}(M)$ with the exponent β , we prove in Theorem 5.9:

THEOREM 1.2. *Let $\Phi : (1, \infty) \rightarrow (0, \infty)$ be a non-increasing function and let*

$$\Psi : \left\{ (\gamma, \beta) \in [0, \infty) \times (1, \infty) : 0 \leq \gamma < \Phi(\beta) < \infty \right\} \rightarrow [0, \infty)$$

be right-continuous in its first argument and such that

$$\Psi(\gamma_1, \beta) \leq \Psi(\gamma_2, \beta) \quad \text{for } 0 \leq \gamma_1 \leq \gamma_2 < \Phi(\beta),$$

with the property that $\|M\|_{\text{BMO}} < \Phi(\beta)$ implies $\mathcal{RH}_\beta(\mathcal{E}(M)) \leq \Psi(\|M\|_{\text{BMO}}, \beta)$. Then, for $\text{sl}_N(M) < \Phi(\beta)$ we have that

$$\mathcal{RH}_\beta(\mathcal{E}(M)) \leq [\Psi(\text{sl}_N(M), \beta)]^N.$$

The point of this observation is that we get explicit exponents β and explicit bounds for $\mathcal{RH}_\beta(\mathcal{E}(M))$ in terms of the sliceable numbers $(\text{sl}_N(M))_{N \geq 1}$. This is applied to BMO-martingales obtained by the fractional gradient $|Z|^\theta$ of our BSDE where $\theta \in [0, 1]$ is the parameter from (4) below that describes the degree of the BSDEs of not being Lipschitz in the Z -component ($\theta = 0$ corresponds to the Lipschitz case, $\theta = 1$ to the quadratic case).

Another contribution concerns the generalized Fefferman inequality [29, Lemma 1.6] (see also [4, Theorem 1.1]). We prove with Theorem 5.20 a more abstract version using adapted random measures that yields in Corollary 5.21 to

$$\left\| \int_0^T |A_t B_t| dt \right\|_p \leq \sqrt{2p} \|A\|_{\mathbb{H}_p(S_2)} \|B\|_{\text{BMO}(S_2)}$$

which improves the asymptotic behavior of the constant from p in [29] to \sqrt{p} .

Chapter 6 The decoupling method for BSDEs originates from [35], where the terminal condition did depend on finitely many increments of a forward diffusion and the generator was Markovian and Lipschitz. The aim of this part of the notes is the further development of this method. Based on observation (3) we first decouple the BSDE (1) in order to get new BSDE

$$Y_t^{(a,b]} = \xi^{(a,b]} + \int_t^T f^{(a,b]}(s, Y_s^{(a,b]}, Z_s^{(a,b]}) ds - \int_t^T Z_s^{(a,b]} dW_s^{(a,b]}$$

and aim to use *a priori estimates* for BSDEs to estimate $\left\| \sup_{s \in [t, T]} |Y_s^{(a,b]} - Y_s| \right\|_p$ and $\left\| \left(\int_t^T |Z_s^{(a,b]} - Z_s|^2 ds \right)^{\frac{1}{2}} \right\|_p$ from above by moments of

$$\xi - \xi^{(a,b]} \quad \text{and} \quad f - f^{(a,b]}.$$

Here we consider generators $f : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ such that $(t, \omega) \mapsto f(t, \omega, y, z)$ is predictable for all (y, z) and there are $L_Y, L_Z \geq 0$ and $\theta \in [0, 1]$ such that

$$(4) \quad |f(t, \omega, y_0, z_0) - f(t, \omega, y_1, z_1)| \leq L_Y |y_0 - y_1| + L_Z [1 + |z_0| + |z_1|]^\theta |z_0 - z_1|$$

for all $(t, \omega, y_0, y_1, z_0, z_1)$. Here $\theta = 0$ represents the Lipschitz case, $\theta = 1$ the quadratic case, and $\theta \in (0, 1)$ the sub-quadratic case. The basic stability result is Theorem 6.4, a special case is:

THEOREM 1.3. *Assume for the BSDE*

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad t \in [0, T],$$

conditions (B1)-(B4) of Chapter 6 for $\theta \in [0, 1]$ and that there is a non-increasing sequence $s = (s_N)_{N \geq 1} \subseteq [0, \infty)$ which dominates the sliceable numbers of the fractional gradient, i.e. $\text{sl}_N^{S_2}(|Z|^\theta) \leq s_N$. Suppose that conditions (B5)-(B6) of Chapter 6 are satisfied for $p \in [2, \infty)$ where in the case $\lim_N s_N > 0$ we additionally assume that $p > p_0(L_Z, \lim_N s_N)$. Then, one has for all $t \in [0, T]$ and $0 \leq a < b \leq T$ that

$$\begin{aligned} & \left\| \sup_{s \in [t, T]} |Y_s^{(a,b]} - Y_s| \right\|_p + \left\| \left(\int_t^T |Z_s^{(a,b]} - Z_s|^2 ds \right)^{\frac{1}{2}} \right\|_p \\ & \leq c \left[\|\xi^{(a,b]} - \xi\|_p + \left\| \int_t^T |f^{(a,b]}(s, Y_s, Z_s) - f(s, Y_s, Z_s)| ds \right\|_p \right]. \end{aligned}$$

In order to apply Theorem 1.3 (Theorem 6.4), and because of general interest, we discuss classes of quadratic and sub-quadratic BSDEs such that the assumptions of Theorems 1.3 and 6.4 are satisfied in Section 6.3. In case of sub-quadratic BSDEs we use the following definition:

DEFINITION 1.4.

- (1) We say that a random variable ξ belongs to cExp provided that there are $(\eta, \mu) \in (0, 1) \times (0, \infty)$ such that

$$|\xi|_{\text{cExp}(\eta, \mu)} := \sup_{t \in [0, T]} (T - t)^{\frac{1}{\eta} - 1} \left\| \mathbb{E}(e^{\mu|\xi|} | \mathcal{F}_t) \right\|_{\infty} < \infty.$$

- (2) For a càdlàg process $Y = (Y_t)_{t \in [0, T]}$ we say that $Y \in \text{cExp}$ provided that there are $(\eta, \mu) \in (0, 1) \times (0, \infty)$ such that

$$|Y|_{\text{cExp}(\eta, \mu)} := \sup_{t \in [0, T]} (T - t)^{\frac{1}{\eta} - 1} \left\| \mathbb{E}(e^{\mu \sup_{s \in [t, T]} |Y_s|} | \mathcal{F}_t) \right\|_{\infty} < \infty.$$

In Theorem 6.13 we prove the following statement:

THEOREM 1.5. *Assume (4) for some $\theta \in (0, 1)$, $\sup_{(t, \omega) \in [0, T] \times \Omega} |f(t, \omega, 0, 0)| < \infty$, and that $\xi \in \text{cExp}$. Then there is a unique solution (Y, Z) to the BSDE (1) in the class where $(Y, |Z|) \in \text{cExp} \times \mathbb{H}_2(S_2)$ ¹. Moreover, for this solution we have that*

$$|Z|^\eta \in \text{BMO}(S_2) \quad \text{for all } \eta \in (0, 1).$$

Theorem 1.5 enables us to apply Theorem 1.3, so that a combination with Theorem 1.1 gives in Corollary 6.23:

COROLLARY 1.6. *Assume (4) for some $\theta \in (0, 1)$, $\sup_{(s, \omega) \in [0, T] \times \Omega} |f(s, \omega, 0, 0)| < \infty$, $\xi \in \text{cExp}$, and that (Y, Z) is the unique solution to the BSDE (1) in the sense of Theorem 1.5. Fix $t \in [0, T]$. Then we have*

$$(5) \quad \text{esssup}_{s \in [0, t]} \|D_s Y_t\|_2 \leq c \sup_{(a, b] \subseteq (0, t]} \frac{1}{\sqrt{b - a}} \left[\|\xi - \xi^{(a, b]}\|_2 + \left\| \int_t^T \sup_{y, z} |f(s, y, z) - f^{(a, b]}(s, y, z)| ds \right\|_2 \right]$$

with the convention that the finiteness of the right-hand side first implies $Y_t \in \mathbb{D}_{1,2}$ and then inequality (5).

The assertion of Corollary 1.6 says that we only need to control directional derivatives of the initial data (ξ, f) on the interval $(0, t]$ (because the perturbations of the original Brownian motion W are only performed on $(a, b] \subseteq (0, t]$) to obtain smoothness of Y_t and that the behaviour of (ξ, f) regarding perturbations on $(t, T]$ does not have any impact - in a sense, we have a smoothing effect.

Finally, let us turn to the L_p -variation of a solution (Y, Z) to our BSDE. Our idea is to use adapted time-nets obtained by a quantile method. This idea is made precise by the following two definitions:

¹The spaces are given in Definitions 6.11 and 5.13 below.

DEFINITION 1.7. Let $p \in [1, \infty)$, $A = (A_t)_{t \in [0, T]}$ be a measurable càdlàg process $A : [0, T] \times \Omega \rightarrow \mathbb{R}$, and $C = (C_t)_{t \in [0, T]}$ be a measurable process $C : [0, T] \times \Omega \rightarrow \mathbb{R}^d$, where \mathbb{R}^d is equipped with the euclidean norm. For a deterministic time-net $\tau = (t_i)_{i=0}^n$ with $0 = t_0 \leq t_1 \leq \dots \leq t_n = T$ we let

$$\text{var}_p([A, C]|\tau) := \sup_{i=1, \dots, n} \left\| \sup_{t_{i-1} \leq s \leq t \leq t_i} |A_t - A_s| \right\|_p + \sup_{i=1, \dots, n} \left\| \left(\int_{t_{i-1}}^{t_i} |C_r|^2 dr \right)^{\frac{1}{2}} \right\|_p.$$

DEFINITION 1.8. Letting $\Lambda : [0, T] \rightarrow (0, \infty)$ be Borel measurable and $n \geq 1$, the time net τ_n^Λ consists of $0 = t_0 < \dots < t_n = T$ such that, for all $i = 1, \dots, n$,

$$\int_{t_{i-1}}^{t_i} \Lambda(r) dr = \frac{1}{n} \int_0^T \Lambda(r) dr.$$

Now we obtain as part of Theorem 6.32 the following result:

THEOREM 1.9. Assume (4) for some $\theta \in (0, 1)$, $\sup_{(t, \omega) \in [0, T] \times \Omega} |f(t, \omega, 0, 0)| < \infty$, $\gamma \in [2, \infty)$, $\xi \in \text{cExp}$, and that

$$\|\xi - \xi^{(a, b]}\|_2 + \left\| \int_a^T \sup_{(y, z) \in \mathbb{R}^{d+1}} |f(r, y, z) - f^{(a, b]}(r, y, z)| dr \right\|_2 \leq \left(\int_a^b \Gamma(r) dr \right)^{\frac{1}{\gamma}}$$

for some integrable Borel function $\Gamma : [0, T] \rightarrow [0, \infty)$. Define the weight function

$$\Lambda(r) := 1 + \|f(r, 0, 0)\|_2 + \Gamma(r).$$

Then one has that

$$\sup_{n \geq 1} \sqrt[n]{\text{var}_2([Y, Z]|\tau_n^\Lambda)} < \infty$$

where the solution is taken from Theorem 1.5.

Theorem 1.9 allows us to control to L_2 -variation of non-Markovian BSDEs by adapted time-nets where only the information of the initial data (ξ, f) is used.

1.3. Notation

Given a metric space M , we let $C(M)$ be the space of all continuous real valued mappings on M . For a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ the space of all random variables $X : \Omega \rightarrow \mathbb{R}$, i.e. Borel measurable maps, is denoted by $\mathcal{L}_0(\Omega, \mathcal{F}, \mathbb{P})$ and equipped with the pseudo-metric

$$(6) \quad d_\Omega(X, Y) := \int_\Omega \frac{|X(\omega) - Y(\omega)|}{1 + |X(\omega) - Y(\omega)|} d\mathbb{P}(\omega).$$

The space $\mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$, $p \in (0, \infty)$, consists of all random variables $X : \Omega \rightarrow \mathbb{R}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\|X\|_p := \left(\int_\Omega |X(\omega)|^p d\mathbb{P}(\omega) \right)^{1/p} < \infty$. As usual, for $p = \infty$ we let $\|X\|_\infty := \text{esssup}_{\omega \in \Omega} |X(\omega)| < \infty$ which yields to the space $\mathcal{L}_\infty(\Omega, \mathcal{F}, \mathbb{P})$. By identifying two random variables X and Y on $(\Omega, \mathcal{F}, \mathbb{P})$ when $X = Y$ \mathbb{P} -a.s., we obtain equivalence classes, denoted by $[X]$, the quasi-normed spaces $(L_p(\Omega, \mathcal{F}, \mathbb{P}), \|\cdot\|_p)$ for $p \in (0, \infty]$, and the complete metric space $(L_0(\Omega, \mathcal{F}, \mathbb{P}), d_\Omega)$ with

$$(7) \quad d_\Omega([X], [Y]) := d_\Omega(X, Y).$$

In Chapters 2 and 3 we carefully distinguish between equivalence classes and random variables, in the later chapters we follow the standard way to identify equivalence classes and random variables if there is no risk of confusion. For two real valued random variables X and Y or \mathbb{R}^n -valued random vectors (X_1, \dots, X_n) and (Y_1, \dots, Y_n) the notation $X \stackrel{d}{=} Y$ and $(X_1, \dots, X_n) \stackrel{d}{=} (Y_1, \dots, Y_n)$ means equality in distribution. We shall use the Burkholder-Davis-Gundy inequalities for continuous local martingales [69, IV.4.1] with $\beta_p \geq 1$ as constant, i.e. given $p \in [1, \infty)$ and a càdlàg real-valued martingale $(M_t)_{t \in [0, T]}$ vanishing at zero, we have

$$(8) \quad \frac{1}{\beta_p} \| [M, M]_T^{\frac{1}{2}} \|_p \leq \| \sup_{t \in [0, T]} |M_t| \|_p \leq \beta_p \| [M, M]_T^{\frac{1}{2}} \|_p$$

where $\beta_p \geq 1$ is an absolute constant and $[M, M]_T$ the quadratic variation of M at time T . We do not need the particular behaviour of the constants β_p , so that we use for the upper and lower bound the same constant. As conventions we use $0^0 := 1$ and

$$A \sim_c B \quad \text{for} \quad \frac{1}{c} A \leq B \leq cA$$

when $A, B \geq 0$ and $c \geq 1$.

CHAPTER 2

A General Factorization

There exist several factorization techniques for random variables and stochastic processes that have the idea to factor a random variable or process through a canonical space that carries the typical information about the problem one is interested in. We will use this idea as an intermediate step to decouple in Chapter 4 the Wiener space and to generate anisotropic Besov spaces. For the Wiener space there are two natural choices as a canonical space: The function space of continuous functions that yields to the Wiener measure and the sequence space $\mathbb{R}^{\mathbb{N}}$ with $\mathbb{N} = \{0, 1, 2, \dots\}$ that yields to an infinite product of standard Gaussian measures. We use the second approach as in [54] and [45], and extend this approach so that no particular distribution (like the Gaussian distribution) is needed and so that it includes the handling of the stochastic processes we work with later. As we work with stochastic processes in continuous time we have to pay attention to the handling of null-sets.

Our factorization procedure yields to the operators \mathcal{C}_S^M that are defined in two steps. Section 2.1 introduces the operators \mathcal{C}_S acting on random variables, in Section 2.2 we extend them to the operators \mathcal{C}_S^M acting on random continuous functions defined on complete σ -compact metric spaces.

2.1. The operators \mathcal{C}_S

We shall work with two complete probability spaces $(\Omega^i, \mathcal{F}^i, \mathbb{P}^i)$, $i = 0, 1$, and random variables $(\xi_k^i)_{k \in I}$, $\xi_k^i : \Omega^i \rightarrow \mathbb{R}$, where $I = \{0, \dots, K\}$ or $I = \{0, 1, 2, \dots\}$, and assume that

- (C1) $\mathcal{F}^{\xi, i} := \sigma(\xi_k^i : k \in I)$,
- (C2) $\mathcal{F}^i = \mathcal{F}^{\xi, i} \vee \mathcal{N}^i$, where $\mathcal{N}^i := \{A^i \in \mathcal{F}^i : \mathbb{P}^i(A^i) = 0\}$,
- (C3) $(\xi_k^0)_{k \in I}$ and $(\xi_k^1)_{k \in I}$ have the same finite-dimensional distributions.

If we omit the superscript i in $\Omega^i, \mathcal{F}^i, \mathbb{P}^i, (\xi_k^i)_{k \in I}, \mathcal{F}^{\xi, i}$, then we consider one of the both probability spaces together with the corresponding random variables and operators introduced later. For the above index set I we let

$$J : \Omega \rightarrow \mathbb{R}^I \quad \text{be given by} \quad J(\omega) := (\xi_k(\omega))_{k \in I}.$$

Let $\mathcal{B}(\mathbb{R}^I)$ be the σ -algebra generated by the cylinder sets, λ be the Lebesgue measure, and $I_T : [0, T] \rightarrow [0, T]$ be the identity, where the time horizon $T > 0$ is fixed in the sequel. Then the following notation will be used:

$\Omega_0 := \Omega$	$\Omega_T := [0, T] \times \Omega$
$\Sigma_0^\xi := \mathcal{F}^\xi$	$\Sigma_T^\xi := \mathcal{B}([0, T]) \otimes \mathcal{F}^\xi$
$\Sigma_0 := \mathcal{F}$	$\Sigma_T := \mathcal{B}([0, T]) \otimes \mathcal{F}$
$\mathbb{P}_0 := \mathbb{P}$	$\mathbb{P}_T := (\lambda \times \mathbb{P})/T$
$\Omega_0^c := \mathbb{R}^I$	$\Omega_T^c := [0, T] \times \mathbb{R}^I$
$\Sigma_0^c := \mathcal{B}(\mathbb{R}^I)$	$\Sigma_T^c := \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}^I)$
$J_0 := J : \Omega_0 \rightarrow \Omega_0^c$	$J_T := (I_T, J) : \Omega_T \rightarrow \Omega_T^c$
$\mathbb{P}_0^c := \text{law}(J_0)$	$\mathbb{P}_T^c := \text{law}(J_T)$

In the table above the letter 'c' stands for *canonical*. When possible, we consider the case of random variables and stochastic processes parallel by the parameter $S \in \{0, T\}$. By the above definitions the probability space $(\Omega_S, \Sigma_S, \mathbb{P}_S)$ is mapped by J_S into the canonical space $(\Omega_S^c, \Sigma_S^c, \mathbb{P}_S^c)$. We also use the following

CONVENTION 2.1. Given a sub- σ -algebra $\mathcal{G}_S \subseteq \Sigma_S$, we will interpret $L_0(\Omega_S, \mathcal{G}_S, \mathbb{P}_S)$ as the space of equivalence classes $[X] \in L_0(\Omega_S, \Sigma_S, \mathbb{P}_S)$ that contain a \mathcal{G}_S -measurable representative.

To keep the notation light we abbreviate, for example, $L_p(\Omega_S) = L_p(\Omega_S, \Sigma_S, \mathbb{P}_S)$ and add only parts to the notation that differ from this convention. Now we start to construct the operators \mathcal{C}_0 and \mathcal{C}_T .

LEMMA 2.2. *For $S \in \{0, T\}$ and any Σ_S -measurable random variable $X : \Omega_S \rightarrow \mathbb{R}$ there is a Σ_S^ξ -measurable random variable $X^\xi : \Omega_S \rightarrow \mathbb{R}$ with $\mathbb{P}_S(X = X^\xi) = 1$.*

PROOF. The case $S = 0$ is known, the case $S = T$ is straightforward: We show that the \mathbb{P}_T -completion of Σ_T^ξ contains Σ_T . It is sufficient to prove that $A \times B \in \overline{\Sigma_T^\xi}^{\mathbb{P}_T}$ for $A \in \mathcal{B}([0, T])$ and $B \in \mathcal{F}$. We find a $B^\xi \in \mathcal{F}^\xi$ such that $\mathbb{P}(B \Delta B^\xi) = 0$. Hence $(A \times B) \Delta (A \times B^\xi) = A \times (B \Delta B^\xi)$ is of \mathbb{P}_T -measure zero. Because $A \times B^\xi \in \Sigma_T^\xi$ we can conclude the proof. \square

LEMMA 2.3. *For $S \in \{0, T\}$ the following assertions hold true:*

- (1) *For each Σ_S^ξ -measurable random variable $X : \Omega_S \rightarrow \mathbb{R}$ there exists a random variable $\widehat{X} : \Omega_S^c \rightarrow \mathbb{R}$ such that*

$$X : \Omega_S \xrightarrow{J_S} \Omega_S^c \xrightarrow{\widehat{X}} \mathbb{R}.$$

- (2) *For $\Sigma_S^{\xi,0}$ -measurable random variables $X, X' : \Omega_S^0 \rightarrow \mathbb{R}$ with $\mathbb{P}_S^0(X = X') = 1$ one has $\mathbb{P}_S^1(\widehat{X} \circ J_S^1 = \widehat{X'} \circ J_S^1) = 1$ where the factorizations $X = \widehat{X} \circ J_S^0$ and $X' = \widehat{X'} \circ J_S^0$ are obtained by part (1).*

PROOF. (1) Starting with $\mathcal{B}(\mathbb{R}^I)$ and $\mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}^I)$, the maps J and J_T generate the σ -algebras \mathcal{F}^ξ and Σ_T^ξ . Hence we can apply the functional representation from the Factorization Lemma [6, p. 62] and (1) follows.

- (2) The assumption implies by a change of variables $\mathbb{P}_S^c(\widehat{X} = \widehat{X'}) = 1$, and by another change of variables the conclusion of assertion (2). \square

The above lemma enables us to introduce the operator \mathcal{C}_S that maps an equivalence class $[X]$ from $L_0(\Omega_S^0)$ to the equivalence class $[\widehat{X} \circ J_S^1]$ in $L_0(\Omega_S^1)$ so that $[X]$ and $[\widehat{X} \circ J_S^1]$ have the same law.

DEFINITION 2.4. For $S \in \{0, T\}$ we define the map $\mathcal{C}_S : L_0(\Omega_S^0) \rightarrow L_0(\Omega_S^1)$ by

$$\mathcal{C}_S(X) = \mathcal{C}_S([X]) := [\widehat{X} \circ J_S^1],$$

where $X \in [X]$ is a $\Sigma_S^{\xi, 0}$ -measurable representative of $[X]$.

Basic properties of \mathcal{C}_0 and \mathcal{C}_T are summarized in Theorem 2.6 below. For its formulation we need a class of functionals $\Phi : (\mathcal{L}_0([0, T]))^n \rightarrow \mathbb{R}$ that excludes, for example, Dirac functionals of type $\Phi(f) := f(t_0)$ for a fixed $t_0 \in [0, T]$:

DEFINITION 2.5. A functional $\Phi : (\mathcal{L}_0([0, T]))^n \rightarrow \mathbb{R}$ is called *consistent* provided that for all probability spaces (A, \mathcal{A}, μ) and jointly measurable processes $X_1, \dots, X_n : [0, T] \times A \rightarrow \mathbb{R}$ the map $F_X : A \rightarrow \mathbb{R}$ with

$$F_X(\omega) = \Phi(X_1(\cdot, \omega), \dots, X_n(\cdot, \omega))$$

is measurable and $\mu(F_X = F_{X'}) = 1$ whenever $(\lambda \times \mu)(X_i \neq X'_i) = 0$ for $i = 1, \dots, n$.

THEOREM 2.6. For $S \in \{0, T\}$, $X_1, \dots, X_n \in \mathcal{L}_0(\Omega_S^0)$ and $Y_i \in \mathcal{C}_S(X_i)$ the following holds true:

- (0) $\xi_k^1 \in \mathcal{C}_0(\xi_k^0)$ for $k \in I$.
- (1) \mathcal{C}_S is a linear isometry and bijection.
- (2) $(Y_1, \dots, Y_n) \stackrel{d}{=} (X_1, \dots, X_n)$.
- (3) For a Borel function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ one has

$$g(Y_1, \dots, Y_n) \in \mathcal{C}_S(g(X_1, \dots, X_n)).$$

For $X, X_1, \dots, X_n \in \mathcal{L}_0(\Omega_T^0)$ and $Y_i \in \mathcal{C}_T(X_i)$ the following is true:

- (4) If $\Phi : (\mathcal{L}_0([0, T]))^n \rightarrow \mathbb{R}$ is consistent, then

$$\Phi(Y_1, \dots, Y_n) \in \mathcal{C}_0(\Phi(X_1, \dots, X_n)).$$

- (5) If X is $\Sigma_T^{\xi, 0}$ -measurable, then there is a $\Sigma_T^{\xi, 1}$ -measurable $Y \in \mathcal{C}_T(X)$ such that for all $t \in [0, T]$ one has

$$Y(t) \in \mathcal{C}_0(X(t)).$$

- (6) For $Y \in \mathcal{L}_0(\Omega_T^1)$ one has $Y \in \mathcal{C}_T(X)$ provided that there is a null-set $\mathcal{N} \subseteq [0, T]$ such that for all $t \in [0, T] \setminus \mathcal{N}$ one has

$$Y(t) \in \mathcal{C}_0(X(t)).$$

PROOF. (0) follows from the definition of \mathcal{C}_0 .

- (1) LINEARITY: Let $a, b \in \mathbb{R}$ and $X, Y \in \mathcal{L}_0(\Omega_S^0)$, and take $\Sigma_S^{\xi, 0}$ -measurable representatives $X^\xi \in [X]$ and $Y^\xi \in [Y]$. Then $aX^\xi + bY^\xi \in a[X] + b[Y]$. From Lemma 2.3 we get that

$$X^\xi(\eta) = \widehat{X}^\xi \circ J_S^0(\eta) \quad \text{and} \quad Y^\xi(\eta) = \widehat{Y}^\xi \circ J_S^0(\eta),$$

for all $\eta \in \Omega_S^0$. Defining point-wise

$$T := a\widehat{X}^\xi + b\widehat{Y}^\xi,$$

we get that $T : \Omega_S^c \rightarrow \mathbb{R}$ is measurable and

$$T(J_S^0(\eta)) = aX^\xi(\eta) + bY^\xi(\eta) \quad \text{for all } \eta \in \Omega_S^0$$

so that $T(J_S^0) \in [aX^\xi + bY^\xi]$. By definition of \mathcal{C}_S ,

$$T(J_S^1) = a\widehat{X^\xi} \circ J_S^1 + b\widehat{Y^\xi} \circ J_S^1 \in \mathcal{C}_S(aX + bY),$$

but is also an element of $a\mathcal{C}_S(X) + b\mathcal{C}_S(Y)$.

ISOMETRY: Because the laws of J_S^0 and J_S^1 coincide, it follows that X and the representatives of $\mathcal{C}_S(X)$ have the same distribution. As $d_S(X, X') = d_S(X - X', 0)$ the property that \mathcal{C}_S is an isometry follows immediately.

BIJECTION: Since \mathcal{C}_S is an isometry, it is an injection. Now let $Y \in \mathcal{L}_0(\Omega_S^1)$ and take Y^ξ to be a $\Sigma_S^{\xi,1}$ -measurable representative of $[Y]$. Then there is a measurable $\widehat{Y^\xi} : \Omega_S^c \rightarrow \mathbb{R}$ such that

$$Y^\xi(\eta) = \widehat{Y^\xi} \circ J_S^1(\eta) \quad \text{for all } \eta \in \Omega_S^1.$$

Now $\eta \mapsto \widehat{Y^\xi} \circ J_S^0(\eta)$ is $\Sigma_S^{\xi,0}$ -measurable and

$$\mathcal{C}_S([\widehat{Y^\xi} \circ J_S^0]) = [Y].$$

(2) The characteristic functions of (X_1, \dots, X_n) and (Y_1, \dots, Y_n) coincide, because for all $(t_1, \dots, t_n) \in \mathbb{R}^n$ and $Y_k \in \mathcal{C}_S(X_k)$ we have

$$\int_{\Omega_S^1} e^{i \sum_{k=1}^n t_k Y_k} d\mathbb{P}_S^1 = \int_{\Omega_S^0} e^{i \sum_{k=1}^n t_k X_k} d\mathbb{P}_S^0$$

where we used (1) and that \mathcal{C}_S keeps the distribution invariant.

(3) We choose X_1^ξ, \dots, X_n^ξ to be $\Sigma_S^{\xi,0}$ -measurable representatives of the classes $[X_1], \dots, [X_n]$, so that

$$X_i^\xi(\eta) = \widehat{X_i^\xi} \circ J_S^0(\eta)$$

for $i = 1, \dots, n$ and all $\eta \in \Omega_S^0$. Next we define the measurable functional $T_Z : \Omega_S^c \rightarrow \mathbb{R}$ as

$$T_Z(\zeta) := g(\widehat{X_1^\xi}(\zeta), \dots, \widehat{X_n^\xi}(\zeta))$$

so that $T_Z \circ J_S^0 = g(X_1^\xi, \dots, X_n^\xi)$. By definition of \mathcal{C}_S we get that

$$\mathcal{C}_S(g(X_1^\xi, \dots, X_n^\xi)) = [T_Z \circ J_S^1].$$

On the other side, by definition of T_Z we have that

$$T_Z \circ J_S^1 = g(\widehat{X_1^\xi} \circ J_S^1, \dots, \widehat{X_n^\xi} \circ J_S^1),$$

which is \mathbb{P}_S^1 -a.s. the same as $g(Y_1, \dots, Y_n)$, where $Y_i \in \mathcal{C}_S(X_i)$. This concludes the proof.

(4) We choose $\Sigma_T^{\xi,0}$ -measurable representatives $X_i^\xi \in [X_i]$, define $Y_i^\xi := \widehat{X_i^\xi} \circ J_T^1$, and get

$$\begin{aligned} F_{Y^\xi}(\omega^1) &= \Phi(\widehat{X_1^\xi}(\cdot, J^1(\omega^1)), \dots, \widehat{X_N^\xi}(\cdot, J^1(\omega^1))), \\ F_{X^\xi}(\omega^0) &= \Phi(\widehat{X_1^\xi}(\cdot, J^0(\omega^0)), \dots, \widehat{X_N^\xi}(\cdot, J^0(\omega^0))). \end{aligned}$$

Letting $\Psi(\zeta) := \Phi(\widehat{X}_1^\xi(\cdot, \zeta), \dots, \widehat{X}_N^\xi(\cdot, \zeta))$, our assumptions yields to a measurable map and $F_{X^\xi} = \Psi \circ J^0$ and $F_{Y^\xi} = \Psi \circ J^1$. Consequently, $F_{Y^\xi} \in \mathcal{C}_0(F_{X^\xi})$. Finally, our assumption yields that F_{X^ξ} and F_X belong to the same equivalence class, and F_{Y^ξ} and F_Y belong to the same equivalence class, so that the proof is complete.

(5) We have that $X = \widehat{X} \circ J_T^0$ for some \widehat{X} , which implies $X(t) = \widehat{X}(t, J^0)$ for all $t \in [0, T]$, and define $Y := \widehat{X} \circ J_T^1$. By construction this implies that $Y(t) = \widehat{X}(t, J^1)$.

(6) Choose $X^\xi \in [X]$ to be $\Sigma_T^{\xi, 0}$ -measurable and $Y^\xi := \widehat{X}^\xi \circ J_T^1$ so that $Y^\xi(t) \in \mathcal{C}_0(X^\xi(t))$ for all $t \in [0, T]$ and

$$Y^\xi \in \mathcal{C}_T(X).$$

Moreover, $\mathbb{P}^0(X^\xi(t) = X(t)) = 1$ for $t \in [0, T] \setminus \mathcal{N}'$ where $\mathcal{N}' \subseteq [0, T]$ is a null-set. Hence $\mathcal{C}_0(X^\xi(t)) = \mathcal{C}_0(X(t))$ for all $t \in [0, T] \setminus \mathcal{N}'$ and $Y(t) \in \mathcal{C}_0(X^\xi(t))$ for all $t \in [0, T] \setminus (\mathcal{N} \cup \mathcal{N}')$. But this implies $Y(t) = Y^\xi(t)$ \mathbb{P}^1 -a.s. for all $t \in [0, T] \setminus (\mathcal{N} \cup \mathcal{N}')$ and $[Y] = [Y^\xi]$. \square

Let us comment on part (4) of Theorem 2.6:

REMARK 2.7.

- (1) Let $L_0([0, T])$ be equipped with the Borel σ -algebra, obtained by the metric of type (7) from Section 1.3, and assume a $\bigotimes_1^n \mathcal{B}(L_0([0, T]))$ -measurable $\Psi : (L_0([0, T]))^n \rightarrow \mathbb{R}$ such that

$$\Phi(f_1, \dots, f_n) = \Psi([f_1], \dots, [f_n]) \quad \text{for } f_1, \dots, f_n \in \mathcal{L}_0([0, T]).$$

Then Φ is consistent.

In fact, the space $L_0([0, T])$ is separable so that its Borel σ -algebra is generated by the open balls. We equip $\mathcal{L}_0([0, T])$ with the smallest σ -algebra $\mathcal{B}(\mathcal{L}_0([0, T]))$ such that $q : \mathcal{L}_0([0, T]) \rightarrow L_0([0, T])$ with $q(f) := [f]$ is measurable. A measurable process $X : [0, T] \times A \rightarrow \mathbb{R}$ generates a canonical map $\hat{X} : A \rightarrow \mathcal{L}_0([0, T])$ that is measurable because

$$\left\{ \omega \in A : \int_0^T \frac{|X(t, \omega) - f(t)|}{1 + |X(t, \omega) - f(t)|} dt < \varepsilon \right\} \in \mathcal{A}$$

for all $\varepsilon > 0$ and $f \in \mathcal{L}_0([0, T])$. Hence we can finish the proof as the composition of two measurable maps is measurable.

- (2) For a measurable $\phi : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ and $g = (g_1, \dots, g_n) \in (\mathcal{L}_0([0, T]))^n$ we obtain a consistent functional by

$$\Phi(g) := \int_0^T \phi(t, g(t)) \chi_{\{\int_0^T |\phi(t, g(t))| dt < \infty\}} dt.$$

Applying Theorem 2.6(4) to the function $\phi(t, x) := |x|^p \wedge L$ with $L, p \in (0, \infty)$ and $x \in \mathbb{R}^n$, we get that

$$\int_0^T (|Y(t)|^p \wedge L) dt \in \mathcal{C}_0 \left(\int_0^T (|X(t)|^p \wedge L) dt \right)$$

for $X(t) = (X_1(t), \dots, X_n(t))$ and $Y(t) = (Y_1(t), \dots, Y_n(t))$. Assuming that $\int_0^T |X(t, \omega)|^p dt < \infty$ for all $\omega \in \Omega^0$, we have

$$\lim_{N \rightarrow \infty} \left(\int_0^T (|X(t, \omega)|^p \wedge N) dt \right) = \int_0^T |X(t, \omega)|^p dt$$

and that $(\int_0^T (|Y(t)|^p \wedge N) dt)_{N \geq 1}$ is a Cauchy sequence in probability. As this sequence converges for all $\omega \in \Omega^1$ (possibly to infinity) we get that

- (a) $\mathbb{P}^1(\{\omega \in \Omega^1 : \int_0^T |Y(t, \omega)|^p dt < \infty\}) = 1$,
- (b) $\int_0^T |Y(t)|^p \chi_{\{\int_0^T |Y(s)|^p ds < \infty\}} dt \in \mathcal{C}_0 \left(\int_0^T |X(t)|^p dt \right)$.

The next theorem shows how some continuity and measurability properties are transferred by the operators \mathcal{C}_0 and \mathcal{C}_T .

THEOREM 2.8. *For $i = 0, 1$ assume right-continuous filtrations $\mathbb{G}^i = (\mathcal{G}_t^i)_{t \in [0, T]}$ with $\mathcal{G}_t^i \subseteq \mathcal{F}^i$ such that \mathcal{G}_0^i contains all null-sets of \mathcal{F}^i and $\mathcal{C}_0(L_0(\Omega^0, \mathcal{G}_t^0)) \subseteq L_0(\Omega^1, \mathcal{G}_t^1)$ for all $t \in [0, T]$. Then the following assertions are true:*

- (1) *If X is path-wise continuous and \mathbb{G}^0 -adapted, then there exists a path-wise continuous \mathbb{G}^1 -adapted process $Y \in \mathcal{L}_0(\Omega_T^1)$ with*

$$Y(t) \in \mathcal{C}_0(X(t)) \quad \text{for all } t \in [0, T].$$

- (2) *One has $\mathcal{C}_T(L_0(\Omega_T^0, \mathcal{P}_T^0)) \subseteq L_0(\Omega_T^1, \mathcal{P}_T^1)$, where \mathcal{P}_T^i are the predictable σ -algebras generated by the \mathbb{G}^i -adapted processes with paths that are left-continuous and have limits from the right.*

PROOF. (1) Taking $\beta(t) \in \mathcal{C}_0(X(t))$ to be \mathcal{G}_t^1 -measurable, Theorem 2.6(2) implies that $(\beta(t))_{t \in [0, T]}$ and $(X(t))_{t \in [0, T]}$ have the same finite-dimensional distributions. Applying Proposition A.1, we note that $Y(t)$ is defined as the a.s.-limit of β_{t_n} , where we may take $t_n \uparrow t$. By our assumption $\beta_{t_n} \in \mathcal{L}_0(\Omega^1, \mathcal{G}_{t_n}^1)$, so that $Y(t)$ is \mathcal{G}_t^1 -measurable. The facts that $(Y(t))_{t \in [0, T]}$ is continuous and a modification of $(\beta_t)_{t \in [0, T]}$ were proven in Proposition A.1.

(2) Applying [49, p. 133, step (b) of the proof of Lemma 2.4] we can approximate any predictable process $X \in \mathcal{L}_0(\Omega_T^0, \mathcal{P}_T^0)$ by a sequence of continuous adapted processes $X^n \in \mathcal{L}_0(\Omega_T^0, \mathcal{P}_T^0)$ with $d_T^0(X^n, X) \rightarrow_n 0$. Applying step (1), we find continuous adapted processes Y^n such that $\lim_n d_T^1(Y^n, Y) \rightarrow_n 0$ for $Y \in \mathcal{C}_T(X)$. Because of $Y^n \in \mathcal{L}_0(\Omega_T^1, \mathcal{P}_T^1)$ we can choose $Y \in \mathcal{L}_0(\Omega_T^1, \mathcal{P}_T^1)$ as well. \square

2.2. The operators \mathcal{C}_S^M as extension of \mathcal{C}_S

We extend our definition of \mathcal{C}_T to decouple random generators of BSDEs. Let M be a locally σ -compact, complete metric space, i.e. there exist compact subsets $\emptyset \neq K_1 \subseteq K_2 \subseteq \dots$, such that $\overline{K_n} = K_n$ and $M = \bigcup_{n=1}^\infty \overset{\circ}{K_n}$.

First we let $f \in \mathcal{L}_0(\Omega_S; C(M))$ if and only if $f : \Omega_S \times M \rightarrow \mathbb{R}$ is a Carathéodory function, i.e. f satisfies that

- (a) $\eta \rightarrow f(\eta, x)$ is measurable for all $x \in M$,
- (b) $x \rightarrow f(\eta, x)$ is continuous for all $\eta \in \Omega_S$.

Equivalently, one can say that $f : \Omega_S \rightarrow C(M)$ is measurable when $C(M)$ is equipped with the smallest σ -algebra $\mathcal{B}(C(M))$ such that for all $x \in M$ the maps $\delta_x : C(M) \rightarrow \mathbb{R}$ with $\delta_x(f) := f(x)$ are Borel-measurable. The space $L_0(\Omega_S; C(M))$ is the space of equivalence classes with $f \sim g$ if $\mathbb{P}_S(f(x) = g(x), x \in M) = 1$. By the next lemma we extend the operator \mathcal{C}_S to $C(M)$ -valued random variables.

LEMMA 2.9. *For $f \in \mathcal{L}_0(\Omega_S^0; C(M))$ there is a $g \in \mathcal{L}_0(\Omega_S^1; C(M))$ with $g(x) \in \mathcal{C}_S(f(x))$ for all $x \in M$. If g_1 and g_2 satisfy this property, then $g_1 = g_2$ \mathbb{P}_S^1 -a.s.*

PROOF. Theorem 2.6 implies that $(f(x))_{x \in M}$ and $(h(x))_{x \in M}$ have the same finite-dimensional distributions for $h(x) \in \mathcal{C}_S(f(x))$, so that the result follows from Proposition A.1. \square

Now we are ready to introduce the extension \mathcal{C}_S^M of \mathcal{C}_S that maps equivalence classes from $L_0(\Omega_S^0; C(M))$ to $L_0(\Omega_S^1; C(M))$ while keeping the distributional properties of the equivalence classes.

DEFINITION 2.10. For $S \in \{0, T\}$ we let

$$\mathcal{C}_S^M : L_0(\Omega_S^0; C(M)) \rightarrow L_0(\Omega_S^1; C(M))$$

such that $\mathcal{C}_S^M([f])$ is the unique equivalence-class whose representatives g satisfy $g(x) \in \mathcal{C}_S(f(x))$ for all $x \in M$. Moreover, $\mathcal{C}_S^M(f) := \mathcal{C}_S^M([f])$.

We close this section with some properties of \mathcal{C}_S^M where we apply Convention 2.1 to $L_0(\Omega_S^i, \mathcal{G}_S^i; C(M))$, i.e. $[f] \in L_0(\Omega_S^i, \mathcal{G}_S^i; C(M))$ provided that there is a representative f that is $\mathcal{G}_S^i/\mathcal{B}(C(M))$ -measurable.

THEOREM 2.11. *The following assertions hold true:*

- (1) *For $M := \mathbb{R}^d$, $f \in \mathcal{L}_0(\Omega_T^0; C(M))$, $X_1, \dots, X_d \in \mathcal{L}_0(\Omega_T^0)$, $g \in \mathcal{C}_T^M(f)$, and $Y_i \in \mathcal{C}_T(X_i)$ one has that*

$$(g(t, Y(t)))_{t \in [0, T]} \in \mathcal{C}_T((f(t, X(t)))_{t \in [0, T]}).$$

- (2) *Let $S \in \{0, T\}$. If one has that $\mathcal{C}_S(L_0(\Omega_S^0, \mathcal{G}_S^0)) \subseteq L_0(\Omega_S^1, \mathcal{G}_S^1)$ for σ -algebras $\mathcal{G}_S^i \subseteq \Sigma_S^i$, then*

$$\mathcal{C}_S^M(L_0(\Omega_S^0, \mathcal{G}_S^0; C(M))) \subseteq L_0(\Omega_S^1, \mathcal{G}_S^1; C(M)).$$

PROOF. (1a) First note that Lemma A.2 implies

$$(f(t, X(t)))_{t \in [0, T]} \in \mathcal{L}_0(\Omega_T^0) \quad \text{and} \quad (g(t, Y(t)))_{t \in [0, T]} \in \mathcal{L}_0(\Omega_T^1).$$

- (1b) Define for $i = 1, \dots, d$, $n \geq 1$, $a_k \in \mathbb{R}$ and a Borel-measurable partition $\bigcup_{k=0}^n B_k = \mathbb{R}$ with $B_k \neq \emptyset$ the processes

$$\begin{aligned} A_i(t) &:= \sum_{k=0}^n a_k 1_{B_k}(X_i(t)) \quad \text{and} \quad A(t) = (A_1(t), \dots, A_d(t)), \\ D_i(t) &:= \sum_{k=0}^n a_k 1_{B_k}(Y_i(t)) \quad \text{and} \quad D(t) = (D_1(t), \dots, D_d(t)). \end{aligned}$$

By Theorem 2.6 we conclude

$$\mathcal{C}_T((f(t, A(t)))_{t \in [0, T]})$$

$$\begin{aligned}
&= \sum_{n_1, \dots, n_d=0}^n \mathcal{C}_T \left(\left(f(t, a_{n_1}, \dots, a_{n_d}) 1_{B_{n_1} \times \dots \times B_{n_d}}(X_1(t), \dots, X_d(t)) \right)_{t \in [0, T]} \right) \\
&= \sum_{n_1, \dots, n_d=0}^n \mathcal{C}_T \left(\left(f(t, a_{n_1}, \dots, a_{n_d}) \right)_{t \in [0, T]} \right) \\
&\quad \mathcal{C}_T \left(\left(1_{B_{n_1} \times \dots \times B_{n_d}}(X_1(t), \dots, X_d(t)) \right)_{t \in [0, T]} \right) \\
&\ni \sum_{n_1, \dots, n_d=0}^n (g(t, a_{n_1}, \dots, a_{n_d}))_{t \in [0, T]} \\
&\quad (1_{B_{n_1} \times \dots \times B_{n_d}}(Y_1(t), \dots, Y_d(t)))_{t \in [0, T]} \\
&= (g(t, D(t)))_{t \in [0, T]},
\end{aligned}$$

where the multiplication of equivalence classes is defined as usual.

(1c) For $L_n(x) := \sum_{k=-4^n}^{4^n-1} \frac{k}{2^n} 1_{[\frac{k}{2^n}, \frac{k+1}{2^n})}(x)$ with $x \in \mathbb{R}$ we let

$$A_i^n(t) := L_n(X_i(t)) \quad \text{and} \quad D_i^n(t) := L_n(Y_i(t))$$

so that $d_T^0(A_i^n, X_i) \rightarrow_n 0$ for $i = 1, \dots, d$. Theorem 2.6 yields $D_i^n \in \mathcal{C}_T(A_i^n)$ and $d_T^1(D_i^n, Y_i) = d_T^1(\mathcal{C}_T(A_i^n), \mathcal{C}_T(X_i)) \rightarrow_n 0$. Because of step (b) and because \mathcal{C}_T is an isometry we obtain the estimates

$$\begin{aligned}
&d_T^1(\mathcal{C}_T((f(t, X(t)))_{t \in [0, T]}), [(g(t, Y(t)))_{t \in [0, T]}]) \\
&\leq d_T^1(\mathcal{C}_T((f(t, X(t)))_{t \in [0, T]}), \mathcal{C}_T((f(t, A^n(t)))_{t \in [0, T]})) \\
&\quad + d_T^1(\mathcal{C}_T((f(t, A^n(t)))_{t \in [0, T]}), [(g(t, D^n(t)))_{t \in [0, T]}]) \\
&\quad + d_T^1([(g(t, D^n(t)))_{t \in [0, T]}], [(g(t, Y(t)))_{t \in [0, T]}]) \\
&= d_T^0((f(t, X(t)))_{t \in [0, T]}, (f(t, A^n(t)))_{t \in [0, T]}) \\
&\quad + d_T^1((g(t, D^n(t)))_{t \in [0, T]}, (g(t, Y(t)))_{t \in [0, T]}).
\end{aligned}$$

Because $f(t, A^n(t)) \rightarrow_n f(t, X(t))$ for all $(t, \omega) \in \Omega_T^0$ we have that

$$d_T^0((f(t, X(t)))_{t \in [0, T]}, (f(t, A^n(t)))_{t \in [0, T]}) \rightarrow_n 0.$$

For the last expression we use that $D_i^n \rightarrow_n Y_i$ in probability implies the convergence $(g(t, D^n(t)))_{t \in [0, T]} \rightarrow_n (g(t, Y(t)))_{t \in [0, T]}$ in probability as well.

(2) From Proposition A.1 it follows that the equivalence-class $\mathcal{C}_S^M(f)$ contains a $(\mathcal{G}_S^1, \mathcal{B}(C(M)))$ -measurable representative. \square

CHAPTER 3

Transference of SDEs

In this chapter we apply the method from Chapter 2 to the Wiener space. The main technical result is Theorem 3.3 below and gives a functional map to move a BSDE from one stochastic basis to another one. For this we do not need any uniqueness of the solution of the BSDE that is moved. By using an independent copy of the Wiener space we generate in Chapter 6 below a twisted copy of our BSDE by this procedure. The comparison of the original BSDE with the twisted copy will yield to the notion of anisotropic smoothness. Theorem 3.3 might also be exploited to map a BSDE to the canonical path-space of continuous functions.

3.1. Setting

For $i = 0, 1$ assume complete probability spaces $(\Omega^i, \mathcal{F}^i, \mathbb{P}^i)$ hosting d -dimensional Brownian motions

$$W^i = (W_t^i)_{t \in [0, T]} = ((W_{t,1}^i, \dots, W_{t,d}^i)^\top)_{t \in [0, T]},$$

where all paths are assumed to be continuous and $W_0^i \equiv 0$. Taking the transposed vector means also that the Brownian motion is considered as column vector. Define the filtrations $\mathbb{F}^i = (\mathcal{F}_t^i)_{t \in [0, T]}$ by $\mathcal{F}_t^i := \sigma(W_s^i : s \in [0, t]) \vee \mathcal{N}^i$ with \mathcal{N}^i being the \mathbb{P}^i -nullsets. Replacing \mathcal{F}^i by \mathcal{F}_T^i we will assume that $\mathcal{F}^i = \mathcal{F}_T^i$. Furthermore, we equip $L_2((0, T]; \ell_2^d)$ with the orthonormal basis $(h_k \otimes e_i)_{k=0, i=1}^{\infty, d}$, where $(h_k)_{k=0}^\infty$ are the $L_2((0, T])$ -normalized Haar functions and e_1, \dots, e_d the unit vectors of ℓ_2^d . The corresponding systems $(\xi_k^i)_{k \in I}$ of random variables from Section 2.1 are given by

$$(9) \quad \mathcal{B}^i := \{g_{k,j}^i : k \geq 0, j = 1, \dots, d\} \quad \text{with} \quad g_{k,j}^i := \int_0^T h_k(t) dW_{t,j}^i,$$

where we take as the representative the finite differences of the j -th coordinate of W^i generated by the Haar function h_k . Because all paths of W^i are continuous we have

$$\sigma(W_{t,j}^i : t \in [0, T]; j = 1, \dots, d) = \sigma(g_{k,j}^i : k = 0, 1, 2, \dots \text{ and } j = 1, \dots, d).$$

The predictable σ -algebras on $(\Omega^i, \mathcal{F}^i, \mathbb{P}^i)$ are denoted by \mathcal{P}^i .

3.2. Results

Before we state the main result we need two lemmas.

LEMMA 3.1. *One has $W_{t,j}^1 \in \mathcal{C}_0(W_{t,j}^0)$ for $j = 1, \dots, d$ and $t \in [0, T]$ so that $\mathcal{C}_0(L_0(\Omega^0, \mathcal{F}_t^0, \mathbb{P}^0)) \subseteq L_0(\Omega^1, \mathcal{F}_t^1, \mathbb{P}^1)$ for $t \in [0, T]$.*

PROOF. The construction and Theorem 2.6(0) imply $W_{t,j}^1 \in \mathcal{C}_0(W_{t,j}^0)$ whenever $t = Tk/2^n$ with $n = 0, 1, 2, \dots$ and $k = 0, \dots, 2^n$. For a $t \in (0, T)$ not of this form we find dyadic $t_n \in [0, T]$ with $t_n \rightarrow t$. Hence $W_{t_n,j}^i \rightarrow W_{t,j}^i$ for $i = 0, 1$ in probability

and Theorem 2.6(1) yields $W_{t,j}^1 \in \mathcal{C}_0(W_{t,j}^0)$. The second part of the statement is a consequence of the first one. \square

LEMMA 3.2. Assume that $K^0 \in \mathcal{L}_2(\Omega_T^0, \mathcal{P}^0)$. Then, for all $j = 1, \dots, d$,

$$\int_0^T K_t^1 dW_{t,j}^1 \in \mathcal{C}_0 \left(\int_0^T K_t^0 dW_{t,j}^0 \right),$$

where $K^1 \in \mathcal{C}_T(K^0)$ is any \mathcal{P}^1 -measurable representative.

PROOF. Let $L \geq 1$, $0 = t_0^L < \dots < t_L^L = T$, and $(\varphi_l^{0,L})_{l=1,\dots,L}$ such that $\varphi_l^{0,L} \in \mathcal{L}_2(\Omega^0, \mathcal{F}_{t_{l-1}^L}^0)$, and $K_t^{0,L} := \sum_{l=1}^L \varphi_l^{0,L} 1_{(t_{l-1}^L, t_l^L]}(t)$ such that

$$\mathbb{E}^0 \int_0^T |K_t^0 - K_t^{0,L}|^2 dt \rightarrow 0 \quad \text{as } L \rightarrow \infty,$$

see [49, Lemma 3.2.4]. Using Theorem 2.6 and Lemma 3.1, letting $\varphi_l^{1,L} \in \mathcal{C}_0(\varphi_l^{0,L})$ and $K_t^{1,L} := \sum_{l=1}^L \varphi_l^{1,L} 1_{(t_{l-1}^L, t_l^L]}(t)$, we get

$$\begin{aligned} \mathcal{C}_0 \left(\int_0^T K_t^0 dW_{t,j}^0 \right) &= \lim_{L \rightarrow \infty} \mathcal{C}_0 \left(\sum_{l=1}^L \varphi_l^{0,L} (W_{t_l^L, j}^0 - W_{t_{l-1}^L, j}^0) \right) \\ &\ni \lim_{L \rightarrow \infty} \sum_{l=1}^L \varphi_l^{1,L} (W_{t_l^L, j}^1 - W_{t_{l-1}^L, j}^1) = \int_0^T K_t^1 dW_{t,j}^1 \end{aligned}$$

where the limits are taken in $L_2(\Omega^1)$ and K^1 is a \mathcal{P}^1 -measurable process that satisfies $\mathbb{E}^1 \int_0^T |K_t^1 - K_t^{1,L}|^2 dt \rightarrow_L 0$. Because of Theorem 2.6(6) we have $K^{1,L} \in \mathcal{C}_T(K^{0,L})$ so that $K^1 \in \mathcal{C}_T(K^0)$ as well. \square

For integers $N, d \geq 1$ and $(\Omega, \mathbb{F}, \mathbb{P}, W)$ being one of the quadruples $(\Omega^i, \mathbb{F}^i, \mathbb{P}^i, W^i)$ we consider

$$(10) \quad L_t = \xi + \int_t^T f(s, K_s) ds - \sum_{j=1}^d \int_t^T g_j(s, K_s) dW_{s,j}$$

where

- (S1) $\xi \in \mathcal{L}_0(\Omega)$,
- (S2) f and g_j are $(\mathcal{P}, \mathcal{B}(C(\mathbb{R}^N)))$ -measurable,
- (S3) $L = (L_t)_{t \in [0, T]}$, $L_t : \Omega \rightarrow \mathbb{R}$, is continuous and \mathbb{F} -adapted,
- (S4) $K = (K_t)_{t \in [0, T]}$, $K_t : \Omega \rightarrow \mathbb{R}^N$, is \mathcal{P} -measurable,
- (S5) $\mathbb{E} \int_0^T [|f(t, K_t)| + |g(t, K_t)|^2] dt < \infty$,
- (S6) (ξ, f, g, K, L, W) satisfies (10) for $t \in [0, T]$ \mathbb{P} -a.s.

Our main technical result is:

THEOREM 3.3. Assume that $(\xi^0, f^0, g^0, K^0, L^0, W^0)$ satisfies (S1)-(S6). Let $\xi^1 \in \mathcal{C}_0(\xi^0)$, $f^1 \in \mathcal{C}_T^{\mathbb{R}^N}(f^0)$ and $g_j^1 \in \mathcal{C}_T^{\mathbb{R}^N}(g_j^0)$ be $(\mathcal{P}^1, \mathcal{B}(C(\mathbb{R}^N)))$ -measurable, $L^1 \in \mathcal{C}_0^{[0, T]}(L^0)$ be \mathbb{F}^1 -adapted and $K_l^1 \in \mathcal{C}_T(K_l^0)$ be \mathcal{P}^1 -measurable for $l = 1, \dots, N$. Then $(\xi^1, f^1, g^1, K^1, L^1, W^1)$ satisfies conditions (S1)-(S6).

PROOF. The existence of suitable measurable representatives can be deduced from a combination of Lemma 3.1 and Theorems 2.8 and 2.11. Using Theorem 2.11(1) we have for $\phi \in \{f, g_j\}$ that $(\phi^1(t, K_t^1))_{t \in [0, T]} \in \mathcal{C}_T((\phi^0(t, K_t^0))_{t \in [0, T]})$. Continuing with Remark 2.7(2) yields that condition (S5) is satisfied for $f^1(t, K_t^1)$ and $g^1(t, K_t^1)$. For a fixed $t \in [0, T]$ we have

$$\mathcal{C}_0(L_t^0) = \mathcal{C}_0(\xi^0) + \mathcal{C}_0 \left(\int_t^T f^0(s, K_s^0) ds \right) - \sum_{j=1}^d \mathcal{C}_0 \left(\int_t^T g_j^0(s, K_s^0) dW_{s,j}^0 \right).$$

Using Remark 2.7(2) with $\phi(t, x) = x$, we have

$$\int_t^T f^1(s, K_s^1) ds \in \mathcal{C}_0 \left(\int_t^T f^0(s, K_s^0) ds \right)$$

for all $t \in [0, T]$. Similarly Lemma 3.2 gives, for $t \in [0, T]$,

$$\int_t^T g_j^1(s, K_s^1) dW_{s,j}^1 \in \mathcal{C}_0 \left(\int_t^T g_j^0(s, K_s^0) dW_{s,j}^0 \right). \quad \square$$

Later, in our application we need that certain properties of the generator transfer. For this purpose we use the following

REMARK 3.4. Assume that $h^0 : \Omega_T^0 \rightarrow C(\mathbb{R}^N)$ is $(\mathcal{P}^0, \mathcal{B}(C(\mathbb{R}^N)))$ -measurable and $h^1 \in \mathcal{C}_T^{\mathbb{R}^N}(h^0)$ is $(\mathcal{P}^1, \mathcal{B}(C(\mathbb{R}^N)))$ -measurable. Then the following holds:

- (1) $h^0(\cdot, \cdot, 0) \stackrel{d}{=} h^1(\cdot, \cdot, 0)$ with respect to $\lambda \times \mathbb{P}^0$ and $\lambda \times \mathbb{P}^1$.
- (2) Given a continuous $H : \mathbb{R}^N \times \mathbb{R}^N \rightarrow [0, \infty)$ such that, for all (t, ω^0, x_0, x_1) ,

$$|h^0(t, \omega^0, x_0) - h^0(t, \omega^0, x_1)| \leq H(x_0, x_1),$$

then we can choose h^1 such that, for all (t, ω^1, x_0, x_1) ,

$$|h^1(t, \omega^1, x_0) - h^1(t, \omega^1, x_1)| \leq H(x_0, x_1).$$

PROOF. (1) follows from Definition 2.10 and Theorem 2.6.

- (2) Given $x \in \mathbb{R}^N$, we have by construction $h^1(x) \in \mathcal{C}_T(h^0(x))$, so that

$$(h^1(x_0), h^1(x_1)) \stackrel{d}{=} (h^0(x_0), h^0(x_1)) \quad \text{for all } x_0, x_1 \in \mathbb{R}^N.$$

This implies that

$$\|h^1(x_0) - h^1(x_1)\|_{L_\infty(\Omega_T^1)} = \|h^0(x_0) - h^0(x_1)\|_{L_\infty(\Omega_T^0)} \leq H(x_0, x_1).$$

Hence, letting

$$\begin{aligned} & \Omega_{T,0}^1 \\ &:= \{(t, \omega^1) \in [0, T] \times \Omega^1 : |h^1(t, \omega^1, x_0) - h^1(t, \omega^1, x_1)| \leq H(x_0, x_1) \\ & \quad \text{for all } x_0, x_1 \in \mathbb{R}^N\} \\ &= \{(t, \omega^1) \in [0, T] \times \Omega^1 : |h^1(t, \omega^1, x_0) - h^1(t, \omega^1, x_1)| \leq H(x_0, x_1) \\ & \quad \text{for all } x_0, x_1 \in \mathbb{Q}^N\}, \end{aligned}$$

we have that $\Omega_{T,0}^1 \in \mathcal{P}^1$ and $\mathbb{P}_T^1(\Omega_{T,0}^1) = 1$. Setting

$$\tilde{h}^1 := \chi_{\Omega_{T,0}^1} h^1 \in \mathcal{C}_T^{\mathbb{R}^N}(h^0),$$

we obtain a $(\mathcal{P}^1, \mathcal{B}(C(\mathbb{R}^N)))$ -measurable map as desired. \square

CHAPTER 4

Anisotropic Besov Spaces on the Wiener Space

In this chapter we introduce anisotropic Besov spaces on the Wiener space by the decoupling method from Chapter 2. The spaces are designed such that non-linear conditional expectations, that are generated by BSDEs, map these spaces into itself (see Chapter 6). This fact will provide variational estimates for solutions to BSDEs. Our approach to define anisotropic Besov spaces is very flexible as it allows different types of spaces, including the classical spaces obtained by the real interpolation method.

4.1. Classical Besov spaces on the Wiener space

In this section we introduce the classical Besov spaces on the Wiener space obtained by the real interpolation method. To do so we first recall the real interpolation method.

4.1.1. Real interpolation method. For detailed information about the real interpolation method the reader is referred (for example) to the monographs [8], [9], or [66]. To define the method in the general context, we say that two Banach spaces (E_0, E_1) form a *compatible couple* provided that there is a Hausdorff space X such that E_0 and E_1 are continuously embedded into X .

DEFINITION 4.1. Given a compatible couple (E_0, E_1) of Banach spaces and $x \in E_0 + E_1$ and $t > 0$, we define the K -functional

$$K(x, t; E_0, E_1) := \inf \{ \|x_0\|_{E_0} + t\|x_1\|_{E_1} : x = x_0 + x_1, x_i \in E_i \}.$$

For $\theta \in (0, 1)$ and $q \in [1, \infty]$ we let $(E_0, E_1)_{\theta, q}$ be the real interpolation space of all $x \in E_0 + E_1$ such that

$$\|x\|_{(E_0, E_1)_{\theta, q}} := \|t^{-\theta} K(x, t; E_0, E_1)\|_{L_q((0, \infty), \frac{dt}{t})} < \infty.$$

The parameter θ is the main interpolation index, the parameter q is a fine-tuning parameter. In most of the cases the space X can be taken in a natural way. Assuming for example that E_1 is continuously embedded into E_0 , which is our typical case later, we can take $X = E_0$ itself. Some basic properties of the real method are the following:

PROPOSITION 4.2. *Let (E_0, E_1) be a compatible couple of Banach spaces, $\theta \in (0, 1)$, and $q \in [1, \infty]$. Then one has the following:*

- (1) $(E_0, E_1)_{\theta, q} = (E_1, E_0)_{1-\theta, q}$ with equal norms for $\theta \in (0, 1)$ and $q \in [1, \infty]$.
- (2) $(E_0, E_1)_{\theta, q_0} \subseteq (E_0, E_1)_{\theta, q_1}$ with a continuous embedding for $\theta \in (0, 1)$ and $0 \leq q_0 \leq q_1 \leq \infty$.

(3) If E_1 is continuously embedded into E_0 , then

$$(E_0, E_1)_{\theta_0, q_0} \subseteq (E_0, E_1)_{\theta_1, q_1} \quad \text{for } 0 < \theta_1 < \theta_0 < 1 \quad \text{and } q_0, q_1 \in [1, \infty],$$

where the embedding is continuous.

4.1.2. Besov spaces on the abstract Wiener space. We assume a separable Hilbert space H , a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and an iso-normal family of Gaussian random variables $(g_h)_{h \in H}$, $g_h : \Omega \rightarrow \mathbb{R}$, i.e.

$$\mathbb{E}g_h = 0 \quad \text{and} \quad \mathbb{E}g_h g_k = \langle h, k \rangle \quad \text{for all } h, k \in H.$$

For $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ and $h_1, \dots, h_n \in H$ this implies

$$\alpha_1 g_{h_1} + \dots + \alpha_n g_{h_n} = g_{\alpha_1 h_1 + \dots + \alpha_n h_n} \quad \text{a.s.},$$

that means that $(g_h)_{h \in H}$ is a Gaussian process. W.l.o.g. we may assume that \mathcal{F} is the completion of $\sigma(g_h : h \in H)$. Let $(\mathbf{h}_n)_{n=0}^\infty$ be the normalised *Hermite-polynomials*, i.e. $\mathbf{h}_n : \mathbb{R} \rightarrow \mathbb{R}$ with $\mathbf{h}_0 \equiv 1$ and

$$\mathbf{h}_n(x) := (-1)^n \frac{1}{\sqrt{n!}} e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}} \quad \text{for } n \geq 1.$$

Letting γ_N be the standard Gaussian measure on \mathbb{R}^N , the Hermite polynomials form an orthogonal basis in $L_2(\mathbb{R}, \gamma_1)$. Now we are in a position to define the Wiener chaos:

DEFINITION 4.3. Let $(e_k)_{k \in I} \subseteq H$ be an orthogonal basis of H . Given $n \geq 1$, the space (of equivalence classes)

$$\mathcal{H}_n := \overline{\text{span} \left\{ \prod_{k \in I} \mathbf{h}_{n_k}(g_{e_k}) : \sum_{k \in I} n_k = n \right\}} \subseteq L_2,$$

where the closure is taken in L_2 , is the n -th Wiener chaos. For $n = 0$ we let \mathcal{H}_0 be space of all equivalence classes that contain a constant.

The space \mathcal{H}_n does not depend on the choice of the orthogonal basis $(e_k)_{k \in I} \subseteq H$. Moreover, one has the fundamental *Wiener chaos expansion*

$$L_2(\Omega, \mathcal{F}, \mathbb{P}) = \oplus_{n=0}^\infty \mathcal{H}_n,$$

in particular the spaces \mathcal{H}_n and \mathcal{H}_m are orthogonal for $n \neq m$. Letting

$$J_n : L_2 \rightarrow \mathcal{H}_n \subseteq L_2$$

be the orthogonal projection onto the n -th chaos, we define

$$\mathbb{D}_{1,2} := \left\{ \xi \in L_2 : \|\xi\|_{\mathbb{D}_{1,2}}^2 := \sum_{n=0}^\infty (n+1) \|J_n \xi\|_2^2 < \infty \right\}.$$

As Malliavin derivative we take $D : \mathbb{D}_{1,2} \rightarrow L_2^H$ with

$$D \left(\prod_{k \in I} \mathbf{h}_{n_k}(g_{e_k}) \right) := \sum_{l \in I} \prod_{k \neq l} \mathbf{h}_{n_k}(g_{e_k}) h'_{n_l}(g_{e_l}) e_l.$$

It is known that

$$Df(g_{h_1}, \dots, g_{h_n}) = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(g_{h_1}, \dots, g_{h_n}) h_k$$

for (say) $f \in C_b^\infty(\mathbb{R}^n)$ and $h_1, \dots, h_n \in H$. If $p \in (2, \infty)$, then we let

$$\mathbb{D}_{1,p} := \{f \in \mathbb{D}_{1,2} : \|f\|_{\mathbb{D}_{1,p}}^p := \|f\|_p^p + \|Df\|_{L_p^H}^p < \infty\}$$

which is consistent with the case $p = 2$. Moreover, we set

$$\mathbb{B}_{p,q}^\theta := (L_p, \mathbb{D}_{1,p})_{\theta,q}.$$

In the case $\dim(H) = d$ we identify L_2 with $L_2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \gamma_d)$ and use the family $g_{(\xi_1, \dots, \xi_d)} : \mathbb{R}^d \rightarrow \mathbb{R}$ given by

$$g_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) := \xi_1 x_1 + \dots + \xi_d x_d.$$

We denote these particular Besov spaces by $\mathbb{B}_{p,q}^\theta(\mathbb{R}^d, \gamma_d)$. To motivate the decoupling method and the corresponding Besov spaces introduced in Sections 4.2 and 4.3 below, we start by the following result:

THEOREM 4.4 ([39, Theorem 3.1]). *Let $p \in [2, \infty)$, $\theta \in (0, 1)$, $q \in [1, \infty]$ and $f \in L_p(\mathbb{R}^d, \gamma_d)$. Then*

$$\|f\|_{\mathbb{B}_{p,q}^\theta(\mathbb{R}^d, \gamma_d)} \sim_{c_{(4.4)}} \|f\|_p + \left\| (1-t)^{-\frac{\theta}{2}} \|f(W_1) - \mathbb{E}(f(W_1)|\mathcal{F}_t)\|_p \right\|_{L_q([0,1], \frac{dt}{1-t})}$$

where $c_{(4.4)} \geq 1$ depends uniquely on (p, θ, q) and $(W_t)_{t \in [0,1]}$ is a d -dimensional $(\mathcal{F}_t)_{t \in [0,1]}$ -Brownian motion.

The above theorem generalizes results from [38]. It is not difficult to verify (cf. Lemma 4.20 below) that

$$\|f(W_1) - \mathbb{E}(f(W_1)|\mathcal{F}_t)\|_p \sim_2 \|f(W_1) - f(W_t + [W'_1 - W'_t])\|_p,$$

where $(W'_t)_{t \in [0,1]}$ is an independent copy of $(W_t)_{t \in [0,1]}$. Denoting

$$W_1^{(t,1]} := W_t + [W'_1 - W'_t],$$

we obtain the equivalent formulation

$$(11) \quad \|f\|_{\mathbb{B}_{p,q}^\theta(\mathbb{R}^d, \gamma_d)} \sim \|f\|_p + \left\| (1-t)^{-\frac{\theta}{2}} \|f(W_1) - f(W_1^{(t,1]})\|_p \right\|_{L_q([0,1], \frac{dt}{1-t})},$$

which yields to the Besov spaces \mathbb{B}_p^Φ in Definition 4.12 below.

4.2. Setting

For $d \geq 1$ and $T > 0$ we fix two standard d -dimensional Brownian motions $W = (W_t)_{t \in [0,T]}$ and $W' = (W'_t)_{t \in [0,T]}$, where all paths are assumed to be continuous, $W_0 \equiv 0$, $W'_0 \equiv 0$, defined on complete probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\Omega', \mathcal{F}', \mathbb{P}')$, where \mathcal{F} and \mathcal{F}' are the completions of $\sigma(W_t : t \in [0, T])$ and $\sigma(W'_t : t \in [0, T])$, respectively. We let

$$\overline{\Omega} := \Omega \times \Omega', \quad \overline{\mathbb{P}} := \mathbb{P} \times \mathbb{P}', \quad \overline{\mathcal{F}} := \overline{\mathcal{F} \otimes \mathcal{F}'}^{\overline{\mathbb{P}}}$$

and extend the Brownian motions W and W' canonically to $\Omega \times \Omega'$. Given a measurable function $\varphi : (0, T] \rightarrow [0, 1]$, we let

$$(12) \quad W_t^\varphi := \int_0^t [1 - \varphi(u)^2]^{\frac{1}{2}} dW_u + \int_0^t \varphi(u) dW'_u$$

and assume again continuity for all trajectories and that $W_0^\varphi \equiv 0$. For example, for $0 \leq a < b \leq T$, this definition yields to

$$W_t^{\chi(a,b]} = \begin{cases} W_t & : 0 \leq t \leq a \\ W_a + W'_t - W'_a & : a \leq t \leq b \\ W_a + (W'_b - W'_a) + (W_t - W_b) & : b \leq t \leq T \end{cases} \quad \bar{\mathbb{P}}\text{-a.s.}$$

The process W^φ is a standard Brownian motion and $(\mathcal{F}_t^\varphi)_{t \in [0, T]}$ will denote its $\bar{\mathbb{P}}$ -augmented natural filtration, i.e.

$$\mathcal{F}_t^\varphi := \sigma(W_s^\varphi : s \in [0, t]) \vee \bar{\mathcal{N}},$$

where $\bar{\mathcal{N}}$ are the $\bar{\mathbb{P}}$ -nullsets from $\bar{\mathcal{F}}$. Identifying $a \in [0, 1]$ with the function $\varphi : (0, T] \rightarrow [0, 1]$ that is constant a , we agree to take the versions

$$W^0 = W \quad \text{and} \quad W^1 = W'.$$

To apply the results from Chapter 3 we use the pairing between

$$(\bar{\Omega}, \mathcal{F}^0, \bar{\mathbb{P}}, \mathbb{F}^0, W^0, \mathcal{B}^0) \quad \text{and} \quad (\bar{\Omega}, \mathcal{F}^\varphi, \bar{\mathbb{P}}, \mathbb{F}^\varphi, W^\varphi, \mathcal{B}^\varphi),$$

where $\mathcal{F}^\psi := \mathcal{F}_T^\psi$, $\mathbb{F}^\psi = (\mathcal{F}_t^\psi)_{t \in [0, T]}$, and \mathcal{B}^ψ is defined like in (9) for $\psi \in \{0, \varphi\}$. The corresponding operators \mathcal{C}_S and \mathcal{C}_S^M from Definitions 2.4 and 2.10 are denoted by $\mathcal{C}_S(\varphi)$ and $\mathcal{C}_S^M(\varphi)$, respectively.

CONVENTION 4.5.

- (1) If needed, we extend a random variable $\xi : \Omega \rightarrow \mathbb{R}$ to $\tilde{\xi} : \bar{\Omega} \rightarrow \mathbb{R}$ by $\tilde{\xi}(\omega, \omega') := \xi(\omega)$. The extension $\tilde{\xi}$ is measurable with respect to \mathcal{F}^0 . In this sense we can apply the operator

$$\mathcal{C}_0(\varphi) : L_0(\bar{\Omega}, \mathcal{F}^0) \rightarrow L_0(\bar{\Omega}, \mathcal{F}^\varphi)$$

to ξ . To simplify the notation, $\tilde{\xi}$ will be usually denoted by ξ as well.

- (2) We denote by ξ^φ the elements of $\mathcal{C}_0(\varphi)(\xi)$ for a random variable $\xi : \Omega \rightarrow \mathbb{R}$ and, for $0 \leq a < b \leq T$, by $\xi^{(a,b]}$ the random variable $\xi^{\chi(a,b]}$, i.e.

$$\xi^\varphi \in \mathcal{C}_0(\varphi)(\xi) \quad \text{and} \quad \xi^{(a,b]} := \xi^{\chi(a,b]}.$$

Because of Lemma 3.1 this notation is consistent with the definition from (12).

4.3. Definition of anisotropic Besov spaces

We start by defining the parameter space

$$\mathfrak{D} := \{\psi \in \mathcal{L}_2((0, T]) : 0 \leq \psi \leq 1\}$$

equipped with the pseudo-metric

$$\delta(\varphi, \psi) := \|\varphi - \psi\|_{L_2((0, T])}.$$

To define our Besov spaces we need some preparations.

LEMMA 4.6. For $\varphi, \psi \in \mathfrak{D}$, $k \geq 0$, and $i \in \{1, \dots, d\}$ one has that

$$\mathbb{E}|g_{k,i}^\varphi - g_{k,i}^\psi|^2 \leq 2\|h_k\|_\infty^2 \int_0^T |\varphi(t)^2 - \psi(t)^2| dt.$$

Consequently, if $\varphi_n, \varphi \in \mathfrak{D}$ are such that $\lim_n \delta(\varphi_n, \varphi) = 0$, then

$$\lim_n \mathbb{E}|g_{k,i}^{\varphi_n} - g_{k,i}^\varphi|^2 = 0.$$

PROOF. (a) Starting from the corresponding definitions we get

$$\begin{aligned} & \mathbb{E}|g_{k,i}^\varphi - g_{k,i}^\psi|^2 \\ &= \mathbb{E} \left| \int_0^T h_k(t) \langle e_i, dW_t^\varphi \rangle - \int_0^T h_k(t) \langle e_i, dW_t^\psi \rangle \right|^2 \\ &= \mathbb{E} \left| \int_0^T h_k(t) \sqrt{1 - \varphi(t)^2} \langle e_i, dW_t \rangle + \int_0^T h_k(t) \varphi(t) \langle e_i, dW_t' \rangle \right. \\ & \quad \left. - \int_0^T h_k(t) \sqrt{1 - \psi(t)^2} \langle e_i, dW_t \rangle - \int_0^T h_k(t) \psi(t) \langle e_i, dW_t' \rangle \right|^2 \\ &= \int_0^T h_k(t)^2 \left[\sqrt{1 - \varphi(t)^2} - \sqrt{1 - \psi(t)^2} \right]^2 dt \\ & \quad + \int_0^T h_k(t)^2 [\varphi(t) - \psi(t)]^2 dt \\ &\leq \int_0^T h_k(t)^2 |\varphi(t)^2 - \psi(t)^2| dt + \int_0^T h_k(t)^2 [\varphi(t) - \psi(t)]^2 dt. \end{aligned}$$

(b) If we assume that $\lim_n \delta(\varphi_n, \varphi) = 0$, then $\varphi_n \rightarrow \varphi$ in probability with respect to the normalized Lebesgue measure on $[0, T]$ and therefore $|\varphi_n^2 - \varphi^2| \rightarrow 0$ in probability as well. The boundedness $|\varphi_n(t)| \leq 1$ and $|\varphi(t)| \leq 1$ yields to $\lim_n \int_0^T |\varphi(t)^2 - \varphi_n(t)^2| dt = 0$ and we can apply part (a). \square

LEMMA 4.7. Let $p \in (0, \infty)$ and $\xi \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$. Then $\delta(\varphi_n, \varphi) \rightarrow_n 0$ implies that

$$\lim_n \|\xi^{\varphi_n} - \xi^\varphi\|_p = 0.$$

PROOF. (a) Assume that ξ is bounded. Given $\varepsilon > 0$ we find $N \geq 1$, $f \in C_b(\mathbb{R}^N)$, and $(\gamma_i)_{i=1}^N \subset \mathcal{B}^0$ such that $\|\xi - f(\gamma_1, \dots, \gamma_N)\|_p < \varepsilon$. Then, by Theorem 2.6,

$$\begin{aligned} & \frac{1}{c_p} \|\xi^{\varphi_n} - \xi^\varphi\|_p \\ &\leq \|\xi^{\varphi_n} - f(\gamma_1^{\varphi_n}, \dots, \gamma_N^{\varphi_n})\|_p + \|f(\gamma_1^{\varphi_n}, \dots, \gamma_N^{\varphi_n}) - f(\gamma_1^\varphi, \dots, \gamma_N^\varphi)\|_p \\ & \quad + \|f(\gamma_1^\varphi, \dots, \gamma_N^\varphi) - \xi^\varphi\|_p \\ &\leq 2\varepsilon + \|f(\gamma_1^{\varphi_n}, \dots, \gamma_N^{\varphi_n}) - f(\gamma_1^\varphi, \dots, \gamma_N^\varphi)\|_p. \end{aligned}$$

We can conclude by $\lim_n \|f(\gamma_1^{\varphi_n}, \dots, \gamma_N^{\varphi_n}) - f(\gamma_1^\varphi, \dots, \gamma_N^\varphi)\|_p = 0$ which follows by Lemma 4.6.

(b) Assuming a general $\xi \in \mathcal{L}_p$, we let $\xi^L := (-L) \vee \xi \wedge L$ for $L > 0$ and obtain, again by Theorem 2.6,

$$\frac{1}{c_p} \|\xi^{\varphi_n} - \xi^\varphi\|_p$$

$$\begin{aligned}
&\leq \|\xi^{\varphi_n} - (\xi^L)^{\varphi_n}\|_p + \|(\xi^L)^{\varphi_n} - (\xi^L)^\varphi\|_p + \|(\xi^L)^\varphi - \xi^\varphi\|_p \\
&= 2\|\xi - \xi^L\|_p + \|(\xi^L)^{\varphi_n} - (\xi^L)^\varphi\|_p.
\end{aligned}$$

Given $\varepsilon > 0$ we find an $L > 0$ such that $2\|\xi - \xi^L\|_p \leq \varepsilon$, so that

$$\frac{1}{c_p} \limsup_n \|\xi^{\varphi_n} - \xi^\varphi\|_p \leq \varepsilon + \lim_n \|(\xi^L)^{\varphi_n} - (\xi^L)^\varphi\|_p \leq \varepsilon.$$

Because $\varepsilon > 0$ was arbitrary, $\lim_n \|\xi^{\varphi_n} - \xi^\varphi\|_p = 0$. \square

As a trivial by-product we get that $\xi^\varphi = \xi^\psi$ \mathbb{P} -a.s. if $\varphi = \psi$ a.e. Now it is convenient to turn \mathfrak{D} into a complete separable metric space.

DEFINITION 4.8. We define the metric space (Δ, δ) as the equivalence classes of the pseudo-metric space (\mathfrak{D}, δ) with

$$\mathfrak{D} = \{\psi \in \mathcal{L}_2((0, T]) : 0 \leq \psi \leq 1\} \quad \text{and} \quad \delta(\varphi, \psi) = \|\varphi - \psi\|_{L_2((0, T])}.$$

Fixing $p \in (0, \infty)$ and $\xi \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$, we obtain a well-defined map

$$F_{\xi, p} : \Delta \rightarrow [0, \infty) \quad \text{by} \quad \varphi \rightarrow \|\xi - \xi^\varphi\|_p.$$

Directly from Lemma 4.7 we get

LEMMA 4.9. For $p \in (0, \infty)$ and $\xi \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ the map $F_{\xi, p} : \Delta \rightarrow [0, \infty)$ is continuous.

PROOF. For $p \in [1, \infty)$ and $\varphi_n \rightarrow \varphi$ we get that

$$|\|\xi - \xi^{\varphi_n}\|_p - \|\xi - \xi^\varphi\|_p| \leq \|\xi^{\varphi_n} - \xi^\varphi\|_p \rightarrow 0$$

as $n \rightarrow \infty$. In the case $p \in (0, 1)$ we use

$$|\mathbb{E}[\|\xi^{\varphi_n} - \xi\|^p - \|\xi^\varphi - \xi\|^p]| \leq \mathbb{E}|\xi^{\varphi_n} - \xi^\varphi|^p. \quad \square$$

DEFINITION 4.10. Let $C^+(\Delta)$ be the space of all non-negative continuous functions $F : \Delta \rightarrow [0, \infty)$. A functional $\Phi : C^+(\Delta) \rightarrow [0, \infty]$ is called *admissible* provided that

- (A1) $\Phi(F + G) \leq \Phi(F) + \Phi(G)$,
- (A2) $\Phi(\lambda F) = \lambda \Phi(F)$ for $\lambda \geq 0$,
- (A3) $\Phi(F) \leq \Phi(G)$ for $0 \leq F \leq G$,
- (A4) $\Phi(F) \leq \limsup_n \Phi(F_n)$ for $\sup_{\varphi \in \Delta} |F_n(\varphi) - F(\varphi)| \rightarrow_n 0$.

EXAMPLE 4.11. Let $A \subseteq \Delta$ be non-empty and let $\alpha : A \rightarrow (0, \infty)$ be an arbitrary weight. Then the functional

$$\Phi(F) := \sup_{\varphi \in A} \frac{F(\varphi)}{\alpha(\varphi)}$$

is admissible. As (A1)-(A3) are obvious, we only check (A4). From $F(\varphi) \leq \limsup_n F_n(\varphi)$ we complete the proof by

$$\frac{F(\varphi)}{\alpha(\varphi)} \leq \limsup_n \left[\sup_{\psi \in A} \frac{F_n(\psi)}{\alpha(\psi)} \right].$$

DEFINITION 4.12. For $p \in (0, \infty)$, $\xi \in L_p(\Omega)$, and an admissible $\Phi : C^+(\Delta) \rightarrow [0, \infty]$ we let $\xi \in \mathbb{B}_p^\Phi$ provided that $\Phi(\varphi \rightarrow \|\xi - \xi^\varphi\|_p) < \infty$ and set

$$\|\xi\|_{\mathbb{B}_p^\Phi} := \left[\mathbb{E}|\xi|^p + \|\xi\|_{\Phi,p}^p \right]^{\frac{1}{p}} \quad \text{with} \quad \|\xi\|_{\Phi,p} := \Phi(\varphi \rightarrow \|\xi - \xi^\varphi\|_p).$$

PROPOSITION 4.13. For $p \in [1, \infty)$ the space \mathbb{B}_p^Φ is a Banach space.

PROOF. The norm properties can be easily verified, we only verify the completeness. Assume a Cauchy sequence $(\xi_n)_{n \geq 1}$, we obtain by the completeness of L_p a limit $\xi = \lim_n \xi_n$ in L_p . To show that the convergence takes place in \mathbb{B}_p^Φ , let $\varepsilon > 0$ and find $n_\varepsilon \geq 1$ such that for all $m, n \geq n_\varepsilon$ we have that $\|\xi_n - \xi_m\|_p^p + \Phi(F_{\xi_n - \xi_m, p})^p < \varepsilon^p$ with $F_{\xi, p}(\varphi) = \|\xi - \xi^\varphi\|_p$. For all $m, n \geq 1$ we have that

$$\begin{aligned} & |F_{\xi_n - \xi_m, p}(\varphi) - F_{\xi_n - \xi, p}(\varphi)| \\ &= \| \|\xi_n - \xi_m - (\xi_n - \xi_m)^\varphi\|_p - \|\xi_n - \xi - (\xi_n - \xi)^\varphi\|_p \| \\ &\leq \| \xi - \xi_m - (\xi - \xi_m)^\varphi \|_p \\ &\leq 2\|\xi - \xi_m\|_p, \end{aligned}$$

so that assumption (A4) implies for $n \geq n_\varepsilon$ that

$$\|\xi_n - \xi\|_p^p + \Phi(F_{\xi_n - \xi, p})^p \leq \lim_m \|\xi_n - \xi_m\|_p^p + \limsup_m \Phi(F_{\xi_n - \xi_m, p})^p \leq 2\varepsilon^p. \quad \square$$

4.4. Connections to real interpolation

Besov spaces (or fractional order Sobolev spaces) on the Wiener space were studied by various authors, see for example [42] and [73]. In this section we relate our approach to those classical Gaussian Besov spaces obtained by the real interpolation method. There are different ways to introduce fractional smoothness. One basic way consists in using the Ornstein-Uhlenbeck semi-group which yields by the Mehler's formula to a decoupling approach. The other approach uses the speed of convergence of conditional expectations which again yields to a decoupling argument. Although these methods are closely related, so far, they have been handled differently.

Below we indicate both methods, where Definition 4.14 corresponds more to the approach based on conditional expectations and Definition 4.17 to Mehler's formula.

DEFINITION 4.14. For $0 = r_0 < r_1 < \dots < r_L = T$, $\theta_l \in (0, 1)$, $q_l \in [1, \infty]$, and $F \in C^+(\Delta)$ we let

$$\Phi_{r_1, \dots, r_L}^{(\theta_1, q_1), \dots, (\theta_L, q_L)}(F) := \sup_{l=1, \dots, L} \left\| (r_l - t)^{-\theta_l/2} F(\chi_{(t, r_l]}) \right\|_{L_{q_l}([r_{l-1}, r_l], \frac{dt}{r_l - t})}.$$

This functional is admissible:

LEMMA 4.15. The functional $\Phi_{r_1, \dots, r_L}^{(\theta_1, q_1), \dots, (\theta_L, q_L)}$ satisfies the conditions (A1), (A2), (A3), and (A4).

PROOF. First we remark that $F(\chi_{(\cdot, r_l]}) : [r_{l-1}, r_l] \rightarrow [0, \infty)$ is a continuous function as $\chi_{(t_n, r_l]} \rightarrow \chi_{(t, r_l]}$ in Δ if $t_n \rightarrow t$ and as F is continuous.

Because the conditions (A1), (A2), (A3) are obvious, we only check (A4). To do so, we assume $F_n, F : \Delta \rightarrow [0, \infty)$ to be continuous with

$$\sup_{\varphi \in \Delta} |F_n(\varphi) - F(\varphi)| \rightarrow_n 0.$$

Then, by the Fatou property of the spaces L_{q_l} ,

$$\begin{aligned} & \sup_{l=1, \dots, L} \left\| (r_l - t)^{-\theta_l/2} F(\chi_{(t, r_l]}) \right\|_{L_{q_l}([r_{l-1}, r_l], \frac{dt}{r_l - t})} \\ &= \sup_{l=1, \dots, L} \left\| (r_l - t)^{-\theta_l/2} \lim_n F_n(\chi_{(t, r_l]}) \right\|_{L_{q_l}([r_{l-1}, r_l], \frac{dt}{r_l - t})} \\ &= \sup_{l=1, \dots, L} \left\| \lim_n [(r_l - t)^{-\theta_l/2} F_n(\chi_{(t, r_l]})] \right\|_{L_{q_l}([r_{l-1}, r_l], \frac{dt}{r_l - t})} \\ &\leq \sup_{l=1, \dots, L} \liminf_n \left\| [(r_l - t)^{-\theta_l/2} F_n(\chi_{(t, r_l]})] \right\|_{L_{q_l}([r_{l-1}, r_l], \frac{dt}{r_l - t})} \\ &\leq \liminf_n \sup_{l=1, \dots, L} \left\| [(r_l - t)^{-\theta_l/2} F_n(\chi_{(t, r_l]})] \right\|_{L_{q_l}([r_{l-1}, r_l], \frac{dt}{r_l - t})}, \end{aligned}$$

which proves (A4). \square

Directly from (11) we obtain

PROPOSITION 4.16. *For $\theta \in (0, 1)$, $p \in [2, \infty)$, $q \in [1, \infty]$, and a Borel function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ with $\mathbb{E}|f(W_1)|^p < \infty$ one has*

$$f \in \mathbb{B}_{p,q}^\theta(\mathbb{R}^d, \gamma_d) \quad \text{if and only if} \quad f(W_1) \in \mathbb{B}_p^{\Phi_1^{(\theta, q)}}.$$

This can be generalized to anisotropic Gaussian Besov spaces by defining

$$f \in \mathbb{B}_{p; q_1, \dots, q_L}^{\theta_1, \dots, \theta_L}(\mathbb{R}^{Ld}, \gamma_{Ld})$$

if and only if

$$f(W_1, W_2 - W_1, \dots, W_L - W_{L-1}) \in \mathbb{B}_p^{\Phi_{1,2,\dots,L}^{(\theta_1, q_1), \dots, (\theta_L, q_L)}}.$$

Whereas Definition 4.14 allows to exchange the Gaussian structure in an anisotropic manner, we explain below a simple isotropic variant which relates to Mehler's formula:

DEFINITION 4.17. Let $K : [0, 1] \rightarrow \mathbb{R}$ be non-negative and Borel-measurable, $q \in [1, \infty)$, and let $\varphi_r : (0, T] \rightarrow \mathbb{R}$ be given by

$$\varphi_r \equiv r \quad \text{for } r \in [0, 1].$$

Then we define

$$\Phi^{(K, q)}(F) := \left(\int_0^1 K(r) |F(\varphi_r)|^q dr \right)^{\frac{1}{q}}.$$

The definition above means that we use the map $\xi \rightarrow \xi^{\varphi_r}$ that exchanges in an isotropic way the full Brownian motion W by its mixture $\sqrt{1 - r^2}W + rW'$. In the same way as for $\Phi_{r_1, \dots, r_L}^{(\theta_1, q_1), \dots, (\theta_L, q_L)}$ we obtain

LEMMA 4.18. *The functional $\Phi^{(K,q)}$ satisfies the conditions (A1), (A2), (A3), and (A4).*

Using the notation

$$\xi(W) = \xi \quad \text{and} \quad \xi(\sqrt{1-r^2}W + rW') = \xi^{\varphi_r}$$

this yields to the expression

$$\left(\int_0^1 K(r) \|\xi(W) - \xi(\sqrt{1-r^2}W + rW')\|_p^q dr \right)^{\frac{1}{q}}.$$

Using the particular kernel

$$K(r) := \frac{2r}{1-r^2} \left(\ln \frac{1}{1-r^2} \right)^{-1-\frac{\theta q}{2}},$$

for $\theta \in (0, 1)$ and $p = q \in (1, \infty)$ this gives

$$\int_0^\infty t^{-1-\frac{\theta p}{2}} \|\xi(W) - \xi(e^{-\frac{t}{2}}W + \sqrt{1-e^{-t}}W')\|_p^p dt.$$

Spaces based on this type of expression were considered in [42, Remark on p. 428] and identified as interpolation spaces.

4.5. The space $\mathbb{B}_p^{\Phi_2}$

In this section we study the space $\mathbb{B}_p^{\Phi_2}$, where the functional $\Phi_2 : C^+(\Delta) \rightarrow [0, \infty]$ is given by

$$\Phi_2(F) := \sup_{0 \leq s < t \leq T} \frac{F(\chi(s, t])}{\sqrt{t-s}}.$$

To describe these spaces we let, for $p \in (0, \infty)$ and a measurable $\lambda : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ with $\mathbb{E} \int_0^T |\lambda_s|^2 ds < \infty$ (\mathbb{R}^d is equipped with the euclidean norm),

$$\begin{aligned} \|\lambda\|_{L_\infty([0, T]; L_p(\Omega))} &:= \operatorname{esssup}_{s \in [0, T]} \|\lambda_s\|_p, \\ \|\lambda\|_{L_p^*(\Omega; L_2([0, T]))} &:= \sup_{0 \leq a < b \leq T} \left\| \left(\frac{1}{b-a} \int_a^b |\lambda_s|^2 ds \right)^{\frac{1}{2}} \right\|_p. \end{aligned}$$

To shorten the notation we also use $\|\lambda_s\|_p = \|\lambda_s\|_p$. By the Lebesgue differentiation theorem (cf. Lemma A.3 below) one has that

$$\begin{aligned} \|\lambda\|_{L_\infty([0, T]; L_2(\Omega))} &= \|\lambda\|_{L_2^*(\Omega; L_2([0, T]))}, \\ \|\lambda\|_{L_p^*(\Omega; L_2([0, T]))} &\leq \|\lambda\|_{L_\infty([0, T]; L_p(\Omega))} \quad \text{for } 2 \leq p < \infty, \\ \|\lambda\|_{L_\infty([0, T]; L_p(\Omega))} &\leq \|\lambda\|_{L_p^*(\Omega; L_2([0, T]))} \quad \text{for } 0 < p \leq 2. \end{aligned}$$

The next theorem, the main result of this section, is motivated as follows: If $\xi \in \mathbb{B}_2^{\Phi_2}$, then $\xi \in \mathbb{D}_{1,2}$ and the quantity $\|\xi\|_{\Phi_2, p}$ enables us to access the Malliavin derivative of ξ without its explicit computation. As in Corollary 1.6 of Section 1.2 announced, this can be exploited in the context of BSDEs to obtain the differentiability of the Y -process without differentiating the BSDE.

THEOREM 4.19. *One has that $\mathbb{B}_2^{\Phi_2} \subseteq \mathbb{D}_{1,2}$ and the following assertions hold true:*

(1) *For $p \in [2, \infty)$ and $\xi \in \mathbb{D}_{1,2} \cap L_p$ one has*

$$\|\xi\|_{\Phi_2, p} \sim_{c_{(4.19)(1), p}} \|D\xi\|_{L_p^*(\Omega; L_2([0, T]))},$$

where $c_{(4.19)(1), p} \geq 1$ depends on p only.

(2) *For $p \in (1, 2)$ and $\xi \in \mathbb{D}_{1,2}$ one has*

$$\frac{1}{c_{(4.19)(2), p}} \|D\xi\|_{L_\infty([0, T]; L_p(\Omega))} \leq \|\xi\|_{\Phi_2, p} \leq c_{(4.19)(2), p} \|D\xi\|_{L_p^*(\Omega; L_2([0, T]))},$$

where $c_{(4.19)(2), p} \geq 1$ depends on p only.

(3) *There is a $\xi \in \mathbb{D}_{1,2}$ such that for all $p \in [1, \infty)$ one has $\xi \in L_p(\Omega)$, $D\xi \in L_p(\Omega; L_2([0, T]))$, and $\xi \notin \mathbb{B}_2^{\Phi_2}$.*

In the inequalities of the theorem above the expressions might be infinite. For the case $p \in (1, 2)$ the result is still incomplete. However, if one is interested in good moment estimates, then the case $p \in [2, \infty)$ seems to be of more interest than the case $p \in (1, 2)$. To prove Theorem 4.19 we let

$$\mathcal{G}_a^b := \sigma(W_t : t \in [0, a]) \vee \sigma(W_t - W_b : t \in [b, T])$$

for $0 \leq a \leq b \leq T$ considered as σ -algebra in $(\Omega, \mathcal{F}, \mathbb{P})$.

LEMMA 4.20. *For $p \in [1, \infty]$, $\xi = (\xi_1, \dots, \xi_m)$ with $\xi_1, \dots, \xi_m \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$, a norm $\|\cdot\|$ on \mathbb{R}^m , and $0 \leq s < t \leq T$ one has*

$$\frac{1}{2} \left\| \|\xi - \xi^{(s, t]}\| \right\|_p \leq \left\| \|\xi - \mathbb{E}_s^{\mathcal{G}_s^t} \xi\| \right\|_p \leq \left\| \|\xi - \xi^{(s, t]}\| \right\|_p,$$

where in the first and last expression ξ_1, \dots, ξ_m are extended to $\overline{\Omega}$ according to Convention 4.5 and the conditional expectation is taken coordinate-wise.

PROOF. By $p \rightarrow \infty$ it is sufficient to show the assertion for $p \in [1, \infty)$. Assuming $p \in [1, \infty)$ it is sufficient to consider $\xi = (\xi_1, \dots, \xi_m)$ of the form

$$\xi = f(W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}})$$

where $0 \leq t_0 < \dots < t_n \leq T$ and $f : \mathbb{R}^{nd} \rightarrow \mathbb{R}^m$ is continuous and bounded. W.l.o.g. we can assume that s and t belong to the partition points. Then

$$\left\| \|\xi - \xi^{(s, t]}\| \right\|_p \leq \left\| \|\xi - \mathbb{E}_s^{\mathcal{G}_s^t} \xi\| \right\|_p + \left\| \|\widetilde{\mathbb{E}_s^{\mathcal{G}_s^t} \xi} - \xi^{(s, t]}\| \right\|_p = 2 \left\| \|\xi - \mathbb{E}_s^{\mathcal{G}_s^t} \xi\| \right\|_p$$

and

$$\left\| \|\xi - \mathbb{E}_s^{\mathcal{G}_s^t} \xi\| \right\|_p = \left\| \|\xi - \mathbb{E}(\xi^{(s, t]} | \sigma(W_r : r \in [0, T]))\| \right\|_p \leq \left\| \|\xi - \xi^{(s, t]}\| \right\|_p. \quad \square$$

PROOF OF THEOREM 4.19. (3) For $l \geq 1$ we take disjoint intervals $(s_l, t_l] \subseteq (0, T]$ with $t_l - s_l = T4^{-l}$ and $t_l < s_{l+1}$. Define

$$A_l := l \cos(W_{s_l, 1})(W_{t_l, 1} - W_{s_l, 1}) \quad \text{and} \quad \xi := \sum_{l=1}^{\infty} A_l.$$

The sum converges in any L_p , $p \in [1, \infty)$, as

$$\sum_{l=1}^{\infty} \|l \cos(W_{s_l, 1})(W_{t_l, 1} - W_{s_l, 1})\|_p \leq c_p \sum_{l=1}^{\infty} l \sqrt{t_l - s_l} < \infty,$$

where $c_p := \|g\|_p$ with $g \sim N(0, 1)$. Moreover,

$$DA_l = l [\cos(W_{s_l,1})\chi_{(s_l,t_l]} - \sin(W_{s_l,1})(W_{t_l,1} - W_{s_l,1})\chi_{(0,s]}]$$

so that

$$\|DA_l\|_{L_q^{L_2((0,T))}} \leq l [\sqrt{t_l - s_l} + \sqrt{T}c_q\sqrt{t_l - s_l}]$$

for $q \in [2, \infty)$. This implies $\xi \in \mathbb{D}_{1,2}$ and $D\xi \in L_q(\Omega; L_2([0, T]))$. On the other hand,

$$\begin{aligned} & \frac{\|A_L - A_L^{(s_L, t_L)}\|_p}{\sqrt{t_L - s_L}} \\ &= L \frac{\|\cos(W_{s_L,1})(W_{t_L,1} - W_{s_L,1} - (W')_{t_L,1} + (W')_{s_L,1})\|_p}{\sqrt{t_L - s_L}} \\ &\geq L\sqrt{2}c_p \|\cos(W_{s_L,1})\|_p \geq L\kappa_p \end{aligned}$$

where $\kappa_p := \sqrt{2}c_p \inf_{s \in [0, T]} \|\cos(W_{s,1})\|_p > 0$. For each $L \geq 1$ this implies

$$\frac{\|\xi - \xi^{(s_L, t_L)}\|_p}{\sqrt{t_L - s_L}} \geq \frac{\|\sum_{l=1}^L (A_l - A_l^{(s_L, t_L)})\|_p}{\sqrt{t_L - s_L}} = \frac{\|A_L - A_L^{(s_L, t_L)}\|_p}{\sqrt{t_L - s_L}} \geq \kappa_p L$$

and therefore $\xi \notin \mathbb{B}_p^{\Phi_2}$.

(1) and (2) Step (a): We prove $\mathbb{B}_2^{\Phi_2} \subseteq \mathbb{D}_{1,2}$. Let $0 \leq a < b \leq T$, and define for $n \geq 1$ the set

$$D_n(a, b) := \{(t_1, \dots, t_n) \in (0, T]^n : \text{there is a } k \text{ such that } t_k \in (a, b]\}.$$

Assume $\xi \in L_2$ with chaos decomposition

$$\xi = \sum_{n=0}^{\infty} I_n(f_n)$$

with symmetric $f_n : ((0, T] \times \{1, \dots, d\})^n \rightarrow \mathbb{R}$, cf. [59, Example 1.1.2]. By Lemma 4.20 the condition $\xi \in \mathbb{B}_2^{\Phi_2}$ is equivalent to the condition

$$\sum_{n=1}^{\infty} n! \|f_n \chi_{D_n(a,b)}\|_{L_2^n}^2 \leq c^2(b-a)$$

for all $0 \leq a < b \leq T$, where $L_2^n := L_2(((0, T] \times \{1, \dots, d\})^n, \mu^n)$ with $\mu := \lambda \otimes \left(\sum_{i=1}^d \delta_{\{i\}}\right)$ and λ being the Lebesgue measure. For $L \geq 1$, $l = 1, \dots, 2^L$, and $n \geq 1$ let

$$D_n^{l,L} := D_n\left(T\frac{l-1}{2^L}, T\frac{l}{2^L}\right),$$

so that

$$\sum_{n=1}^{\infty} n! \|f_n \chi_{D_n^{l,L}}\|_{L_2^n}^2 \leq c^2 2^{-L}.$$

Summing up over l gives for all $N \geq 1$ that

$$\sum_{n=1}^N n! \sum_{l=1}^{2^L} \|f_n \chi_{D_n^{l,L}}\|_{L_2^n}^2 \leq c^2.$$

Let Δ_n^L be the union of all dyadic half-open cubes

$$\left(T \frac{l_1-1}{2^L}, T \frac{l_1}{2^L}\right] \times \cdots \times \left(T \frac{l_n-1}{2^L}, T \frac{l_n}{2^L}\right]$$

with $l_1, \dots, l_n \in \{1, \dots, 2^L\}$ pair-wise distinct. Then

$$\Delta_n^L \subseteq \bigcup_{l=1}^{2^L} D_n^{l,L}$$

and

$$\text{card}\{l \in \{1, \dots, 2^L\} : (t_1, \dots, t_n) \in D_n^{l,L}\} = n \quad \text{for all } (t_1, \dots, t_n) \in \Delta_n^L.$$

Now we get that

$$\sum_{n=1}^N n! n \|f_n \chi_{\Delta_n^L}\|_{L_2^n}^2 \leq \sum_{n=1}^N n! \sum_{l=1}^{2^L} \|f_n \chi_{D_n^{l,L}}\|_{L_2^n}^2 \leq c^2.$$

By $L \rightarrow \infty$ it follows that

$$\sum_{n=1}^N n! n \|f_n\|_{L_2^n}^2 \leq c^2.$$

Finally, $N \rightarrow \infty$ gives $\xi \in \mathbb{D}_{1,2}$.

Step (b): Let $\xi \in \mathbb{D}_{1,2}$ with chaos expansion $\xi = \sum_{n=0}^{\infty} I_n(f_n)$ obtained by symmetric f_n and fix $b \in (0, T]$. Consider the processes $(\mu_t^b(i))_{t \in [0, b]}$ from Lemma A.10, so that for $p \in (1, \infty)$ and $a \in [0, b)$ we have that

$$(13) \quad \left\| \xi - \xi^{(a,b)} \right\|_p \sim_2 \left\| \xi - \mathbb{E}(\xi | \mathcal{G}_a^b) \right\|_p \sim_{c_{(A.10)}} \left\| \left(\int_a^b |\mu_s^b|^2 ds \right)^{\frac{1}{2}} \right\|_p$$

where Lemma 4.20 is exploited in the first equivalence. For $s \in [a, b]$ and $n \geq 0$ let $t_k^n := a + (k/2^n)(b-a)$ for $k = 0, \dots, 2^n$ and

$$b_n(s) := \inf \{t_k^n : s \leq t_k^n, k = 0, \dots, 2^n\}.$$

Using (13) and Lemma 4.20 we get that

$$\begin{aligned} c_{(A.10)} \sqrt{b-a} \|\xi\|_{\Phi_{2,p}} &\geq c_{(A.10)} \left(\sum_{k=1}^{2^n} \left\| \xi - \mathbb{E}(\xi | \mathcal{G}_{t_{k-1}^n}^{t_k^n}) \right\|_p^2 \right)^{\frac{1}{2}} \\ &\geq \left(\sum_{k=1}^{2^n} \left\| \left(\int_{(t_{k-1}^n, t_k^n]} |\mu_s^{b_n(s)}|^2 ds \right)^{\frac{1}{2}} \right\|_p^2 \right)^{\frac{1}{2}}. \end{aligned}$$

For $p \in [2, \infty)$ we continue by Fatou's lemma to

$$\begin{aligned} c_{(A.10)} \sqrt{b-a} \|\xi\|_{\Phi_{2,p}} &\geq \liminf_n \left(\sum_{k=1}^{2^n} \left\| \left(\int_{(t_{k-1}^n, t_k^n]} |\mu_s^{b_n(s)}|^2 ds \right)^{\frac{1}{2}} \right\|_p^2 \right)^{\frac{1}{2}} \\ &\geq \liminf_n \left\| \left(\sum_{k=1}^{2^n} \int_{(t_{k-1}^n, t_k^n]} |\mu_s^{b_n(s)}|^2 ds \right)^{\frac{1}{2}} \right\|_p \end{aligned}$$

$$\begin{aligned}
&= \liminf_n \left\| \left(\int_{(a,b]} |\mu_s^{b_n(s)}|^2 ds \right)^{\frac{1}{2}} \right\|_p \\
&\geq \left\| \liminf_n \left(\int_{(a,b]} |\mu_s^{b_n(s)}|^2 ds \right)^{\frac{1}{2}} \right\|_p \\
&\geq \left\| \left(\int_{(a,b]} \liminf_n |\mu_s^{b_n(s)}|^2 ds \right)^{\frac{1}{2}} \right\|_p.
\end{aligned}$$

For $p \in (1, 2)$ we get

$$\begin{aligned}
&c(A.10) \sqrt{b-a} \|\xi\|_{\Phi_{2,p}} \\
&\geq \liminf_n \left(\sum_{k=1}^{2^n} \left\| \left(\int_{(t_{k-1}^n, t_k^n]} |\mu_s^{b_n(s)}|^2 ds \right)^{\frac{1}{2}} \right\|_p^2 \right)^{\frac{1}{2}} \\
&\geq \liminf_n \left(\int_{(a,b]} \|\mu_s^{b_n(s)}\|_p^2 ds \right)^{\frac{1}{2}} \\
&\geq \left(\int_{(a,b]} \|\liminf_n |\mu_s^{b_n(s)}|\|_p^2 ds \right)^{\frac{1}{2}}.
\end{aligned}$$

Summarizing, this yields to

$$(14) \quad \|\xi\|_{\Phi_{2,p}} \geq \frac{1}{c(A.10)} \begin{cases} \left\| \left(\frac{1}{b-a} \int_{(a,b]} \liminf_n |\mu_s^{b_n(s)}|^2 ds \right)^{\frac{1}{2}} \right\|_p & : p \in [2, \infty) \\ \left(\frac{1}{b-a} \int_{(a,b]} \|\liminf_n |\mu_s^{b_n(s)}|\|_p^2 ds \right)^{\frac{1}{2}} & : p \in (1, 2) \end{cases}.$$

Now we observe that

$$\begin{aligned}
&\lim_n \int_{(a,b]} \mathbb{E} |\mu_s^{b_n(s)}(i) - D(s, i) \xi|^2 ds \\
&= \lim_n \int_{(a,b]} \sum_{k=1}^{\infty} k^2 (k-1)! \|f_k((s, i), \cdot)\|_{L_2^{k-1}}^2 (\chi_{((0,s] \cup (b_n(s), T])^{k-1}} - 1) ds \\
&= 0
\end{aligned}$$

which follows by dominated convergence since

$$\begin{aligned}
&\int_{(a,b]} \sum_{k=1}^{\infty} k^2 (k-1)! \|f_k((s, i), \cdot)\|_{L_2^{k-1}}^2 (\chi_{((0,s] \cup (b_n(s), T])^{k-1}} - 1) ds \\
&\leq \int_{(0,1]} \sum_{k=1}^{\infty} k^2 (k-1)! \|f_k((s, i), \cdot)\|_{L_2^{k-1}}^2 ds \\
&\leq \|\xi\|_{\mathbb{D}_{1,2}}^2.
\end{aligned}$$

Hence there is a sub-sequence $(n_l)_{l=1}^{\infty}$ such that $\lim_l \mu_s^{b_{n_l}(s)} = D(s, \cdot) \xi \lambda \otimes \mathbb{P}$ a.e. on $(a, b] \times \Omega$. Observing that (14) holds for the sub-sequence $(n_l)_{l=1}^{\infty}$ without modification as well, the desired lower bounds of $\|\xi\|_{\Phi_{2,p}}$ follow.

Step (c): We verify the upper bounds of (1) and (2). Let us first assume that ξ is smooth like in Proposition A.4, i.e. by using the Haar system as orthogonal basis we may assume that

$$\xi = f(W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}}),$$

where $0 = t_0 < \dots < t_n = T$ and $f \in C^\infty(\mathbb{R}^{nd})$ is bounded with bounded derivatives of all orders (the bounds for the derivatives can depend on their order). By a possible redefinition of f we can assume w.l.o.g. that $a = t_k < t_l = b$. We get

$$D\xi = \sum_{i=1}^n \nabla_i f(W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}}) \chi_{(t_{i-1}, t_i]},$$

where ∇_i is the d -dimensional gradient acting on the i -block of variables. We fix $\xi_1, \dots, \xi_k, \xi_{l+1}, \dots, \xi_n \in \mathbb{R}^d$ and let

$$\begin{aligned} f_\xi(\eta_{k+1}, \dots, \eta_l) &:= f(\xi_1, \dots, \xi_k, \eta_{k+1}, \dots, \eta_l, \xi_{l+1}, \dots, \xi_n), \\ f_\xi^0(\eta_{k+1}, \dots, \eta_l) &:= f_\xi\left(\eta_{k+1} \sqrt{\delta_{k+1}}, \dots, \eta_l \sqrt{\delta_l}\right) \end{aligned}$$

for $\delta_i := t_i - t_{i-1}$. Moreover, we note that

$$\begin{aligned} \|\xi - \mathbb{E}(\xi | \mathcal{G}_a^b)\|_p &= \|f(W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}}) - \\ &\quad \mathbb{E}_{k+1}^l f(W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}})\|_p, \end{aligned}$$

where \mathbb{E}_{k+1}^l is the expected value with respect to the increments

$$(W_{t_{k+1}} - W_{t_k}, \dots, W_{t_l} - W_{t_{l-1}}).$$

Applying Lemma A.7 yields to

$$\begin{aligned} &\|f_\xi(W) - \mathbb{E}f_\xi(W)\|_p \\ &= \left\| f_\xi^0\left(\frac{W_{t_{k+1}} - W_{t_k}}{\sqrt{\delta_{k+1}}}, \dots, \frac{W_{t_l} - W_{t_{l-1}}}{\sqrt{\delta_l}}\right) \right. \\ &\quad \left. - \mathbb{E}f_\xi^0\left(\frac{W_{t_{k+1}} - W_{t_k}}{\sqrt{\delta_{k+1}}}, \dots, \frac{W_{t_l} - W_{t_{l-1}}}{\sqrt{\delta_l}}\right) \right\|_p \\ &\leq c_{(A.7)} \left\| \left(\sum_{i=1}^{l-k} \left| \nabla_i f_\xi^0\left(\frac{W_{t_{k+1}} - W_{t_k}}{\sqrt{\delta_{k+1}}}, \dots, \frac{W_{t_l} - W_{t_{l-1}}}{\sqrt{\delta_l}}\right) \right|^2 \right)^{\frac{1}{2}} \right\|_p \\ &= c_{(A.7)} \left\| \left(\sum_{i=k+1}^l \delta_i |\nabla_i f(\xi_1, \dots, \xi_k, W_{t_{k+1}} - W_{t_k}, \right. \right. \\ &\quad \left. \left. \dots, W_{t_l} - W_{t_{l-1}}, \xi_{l+1}, \dots, \xi_n) |^2 \right)^{\frac{1}{2}} \right\|_p \end{aligned}$$

and

$$\begin{aligned} &\|\xi - \mathbb{E}(\xi | \mathcal{G}_a^b)\|_p \\ &= \|f(W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}}) - \\ &\quad \mathbb{E}_{k+1}^l f(W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}})\|_p \end{aligned}$$

$$\begin{aligned}
&\leq c_{(A.7)} \left\| \left(\sum_{i=k+1}^l \delta_i |\nabla_i f(W)|^2 \right)^{\frac{1}{2}} \right\|_p \\
&= c_{(A.7)} \left\| \left(\sum_{i=k+1}^l \int_{(t_{i-1}, t_i]} |\nabla_i f(W)|^2 ds \right)^{\frac{1}{2}} \right\|_p \\
&= c_{(A.7)} \left\| \left(\int_{(a,b]} |D_s \xi|^2 ds \right)^{\frac{1}{2}} \right\|_p.
\end{aligned}$$

Now we assume the general case and let $q := p \vee 2$. Our assumptions in (1) and (2) and under the assumption that the right-hand sides in (1) and (2) are finite, we have that $\xi \in \mathbb{D}_{1,2} \cap L_q$ and $D\xi \in L_q^{L_2[0,T]}$. Using Proposition A.4 we find smooth ξ_n such that

$$\xi_n \rightarrow \xi \quad \text{in } L_q \quad \text{and} \quad D\xi_n \rightarrow D\xi \quad \text{in } L_q^{L_2[0,T]}.$$

Therefore by approximation,

$$(15) \quad \|\xi - \mathbb{E}(\xi | \mathcal{G}_a^b)\|_p \leq c_{(A.7)} \left\| \left(\int_{(a,b]} |D_s \xi|^2 ds \right)^{\frac{1}{2}} \right\|_p$$

under the assumptions (1) and (2). Dividing by $\sqrt{b-a}$ and taking the supremum over $0 \leq a < b \leq T$ gives the upper bound of $\|\xi\|_{\Phi_2, p}$. \square

4.6. An embedding theorem for functionals of bounded variation

We extend the approach from Section 4.5 to the functionals $\Phi_r : C^+(\Delta) \rightarrow [0, \infty]$, $r \in [2, \infty)$, given by

$$(16) \quad \Phi_r(F) := \sup_{0 \leq s < t \leq T} \frac{F(\chi_{(s,t]})}{(t-s)^{\frac{1}{r}}}.$$

DEFINITION 4.21. A Borel function $g : \mathbb{R} \rightarrow \mathbb{R}$ is of bounded variation provided that

$$V(g) := \sup_{-\infty < x_0 < \dots < x_n < \infty} \sum_{k=1}^n |f(x_k) - f(x_{k-1})| < \infty.$$

It follows from the definition that a function of bounded variation is bounded. A typical example is $g = \chi_{[K, \infty)}$ where $V(\chi_{[K, \infty)}) = 1$. Now we get the following embedding:

THEOREM 4.22. Let $r \in [2, \infty)$, $p \in [1, \infty)$, $\xi \in \mathbb{B}_p^{\Phi_r}$ and $g : \mathbb{R} \rightarrow \mathbb{R}$ be of bounded variation. Assume that the law of ξ has a bounded density ρ . Then, for all $q \in [1, \infty)$,

$$g(\xi) \in \mathbb{B}_q^{\Phi_{\tilde{r}}} \quad \text{with} \quad \tilde{r} := \frac{p+1}{p}qr.$$

PROOF. We use [3, Theorem 2.4] and get that

$$\left(\mathbb{E} |g(\xi) - (g(\xi))^{(s,t]}|^q \right)^{\frac{1}{q}} = \left(\mathbb{E} |g(\xi) - (g(\xi^{(s,t]}))|^q \right)^{\frac{1}{q}}$$

$$\leq 3^{\frac{q+1}{q}} \left(\sup_{x \in \mathbb{R}} \rho(x) \right)^{\frac{1}{q} \frac{p}{p+1}} V(g) \|\xi - \xi^{(s,t)}\|_p^{\frac{1}{q} \frac{p}{p+1}}.$$

Dividing by $(t-s)^{\frac{1}{r} \frac{1}{q} \frac{p}{p+1}}$ gives the assertion. \square

In view of Example 4.24 the following limiting case is important:

COROLLARY 4.23. *If $r \in [2, \infty)$, $\xi \in \bigcap_{p \in [1, \infty)} \mathbb{B}_p^{\Phi_r}$ has a bounded density, and if $g : \mathbb{R} \rightarrow \mathbb{R}$ is of bounded variation, then*

$$g(\xi) \in \bigcap_{q \in [1, \infty)} \bigcap_{\tilde{r} \in (qr, \infty)} \mathbb{B}_q^{\Phi_{\tilde{r}}}.$$

4.7. Examples

4.7.1. Forward diffusions. The Malliavin differentiability of diffusions is well investigated, see for example [59]. So the following is expected:

EXAMPLE 4.24. Let

$$X_t = x_0 + \int_0^t \sigma(s, X_s) dW_s + \int_0^t b(s, X_s) ds$$

where $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d$ and $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ are bounded and continuous, and satisfy

$$|\sigma(t, x) - \sigma(t, y)| + |b(t, x) - b(t, y)| \leq L|x - y| \quad \text{for some } L \geq 0.$$

By the proof of [35, Theorem 3] this implies for $p \in [2, \infty)$ that

$$\|X_T^\varphi - X_T\|_p \leq c \left(\int_0^T \varphi(r)^2 dr \right)^{\frac{1}{2}}$$

with $c = c(p, T, b, \sigma) > 0$. In particular, for $X_T = (X_T^1, \dots, X_T^d)$,

$$X_T^i \in \bigcap_{p \in (0, \infty)} \mathbb{B}_p^{\Phi_2}$$

which follows by using $\varphi = \chi_{(s,t]}$ for $0 \leq s < t \leq T$.

4.7.2. Local time. One can look at the fractional smoothness of local times $(L_t^\alpha)_{t \in (0, T], \alpha \in \mathbb{R}}$ of a one-dimensional Brownian motion from different points of view: In [16, 15] the smoothness with respect to the state variable α is under consideration, whereas in [61, 1] (with a generalization in [74]) the smoothness in ω for fixed (t, α) is investigated within the interpolation spaces generated by the Ornstein-Uhlenbeck operator. Our result complements [1, Theorem 1]. The smoothness obtained in [1] is strictly smaller than $1/2$. In Theorem 4.25 and Corollary 4.27 below we show that in the class of Besov spaces \mathbb{B}_p^Φ the function Φ_r defined in (16) with $r = 4$ is the correct one. Interpreting Φ_2 as smoothness 1, the function Φ_4 corresponds to the smoothness $1/2$. Our approach is similar to [1]: First we investigate the functional N_T^L and then the local time itself by Tanaka's formula.

THEOREM 4.25. Let $d = 1$, $\alpha \in \mathbb{R}$, and

$$N_T^\alpha := \int_{(0,T]} \chi_{\{W_t > \alpha\}} dW_t.$$

Then, for all $p \in (1, \infty)$, one has that

$$N_T^\alpha \in \mathbb{B}_p^{\Phi_4} \setminus \left[\bigcup_{r \in [2,4)} \mathbb{B}_p^{\Phi_r} \right].$$

REMARK 4.26. (1) The natural range for the parameter r in Φ_r is $r \in [2, \infty)$ so that we used the condition $r \in [2, 4)$ instead of the equivalent one $r \in (0, 4)$.

(2) It follows that $N_T^\alpha \in \mathbb{B}_p^{\Phi_4}$ for all $p \in (0, \infty)$, but for the part $N_T^\alpha \notin \bigcup_{r \in [2,4)} \mathbb{B}_p^{\Phi_r}$ our argument uses $p > 1$.

PROOF OF THEOREM 4.25. (a) Denote $\xi = N_T^\alpha$. For the part $N_T^\alpha \in \mathbb{B}_p^{\Phi_4}$ we only need to consider the case $p \in [2, \infty)$ and let $0 \leq a < b \leq T$. Then, a.s.,

$$\begin{aligned} \xi - \xi^{(a,b]} &= \int_{(a,b]} \chi_{\{W_t > \alpha\}} dW_t - \int_{(a,b]} \chi_{\{W_t^{(a,b]} > \alpha\}} dW_t^{(a,b]} \\ &\quad + \int_{(b,T]} [\chi_{\{W_t > \alpha\}} - \chi_{\{W_t^{(a,b]} > \alpha\}}] dW_t \end{aligned}$$

where we use that $(\int_{(0,T]} \chi_{\{W_t > \alpha\}} dW_t)^{(a,b]} = \int_{(0,T]} \chi_{\{W_t^{(a,b]} > \alpha\}} dW_t^{(a,b]}$ a.s. which can be proved by approximating the stochastic integral by Riemann sums that converge in L_2 towards the original integral and to apply the $\cdot^{(ab)}$ -operation to the Riemann sums. Then, by the Burkholder-Davis-Gundy inequalities,

$$\begin{aligned} &\|\xi - \xi^{(a,b]}\|_p \\ &\leq 2 \left\| \int_{(a,b]} \chi_{\{W_t > \alpha\}} dW_t \right\|_p + \left\| \int_{(b,T]} [\chi_{\{W_t > \alpha\}} - \chi_{\{W_t^{(a,b]} > \alpha\}}] dW_t \right\|_p \\ &\leq \beta_p \left[2\sqrt{b-a} + \left\| \left(\int_b^T |\chi_{\{W_t > \alpha\}} - \chi_{\{W_t^{(a,b]} > \alpha\}}|^2 dt \right)^{\frac{1}{2}} \right\|_p \right] \\ &= \beta_p \left[2\sqrt{b-a} + \left\| \int_b^T \chi_{I_\alpha(W_b, W_b^{(a,b]})} (W_t - W_b) dt \right\|_q^{\frac{1}{2}} \right] \end{aligned}$$

for $q := p/2 \in [1, \infty)$ and

$$I_\alpha(u, v) := (\alpha - u, \alpha - v] \cup (\alpha - v, \alpha - u] = (\alpha - \max\{u, v\}, \alpha - \min\{u, v\}].$$

Let $-\infty < A < B < \infty$ and define the function $f_{A,B} : \mathbb{R} \rightarrow \mathbb{R}$ by

$$f_{A,B}(x) := \begin{cases} 0 & : x \leq A \\ (x - A)^2 & : A < x < B \\ (B - A)^2 + 2(B - A)(x - B) & : B \leq x \end{cases}$$

By the Itô-Tanaka formula and the occupation times formula (see [69, VI.1.5 and VI.1.6]) we get that, a.s.,

$$f_{A,B}(W_T - W_b) = f_{A,B}(0) + \int_{(b,T]} f'_{A,B}(W_t - W_b) dW_t + \int_b^T \chi_{(A,B]}(W_t - W_b) dt.$$

This gives that

$$\begin{aligned} & \left\| \int_b^T \chi_{(A,B]}(W_t - W_b) dt \right\|_q \\ & \leq \|f_{A,B}(W_T - W_b) - f_{A,B}(0)\|_q + \left\| \int_{(b,T]} f'_{A,B}(W_t - W_b) dW_t \right\|_q \\ & \leq \|f'_{A,B}\|_\infty \left[\|W_T - W_b\|_q + \beta_q \sqrt{T-b} \right] \\ & = 2(B-A) \left[\|W_T - W_b\|_q + \beta_q \sqrt{T-b} \right] \\ & \leq 4\beta_q \sqrt{T-b} (B-A). \end{aligned}$$

Then

$$\begin{aligned} \left\| \int_b^T \chi_{I_\alpha(W_b, W_b^{(a,b]})}(W_t - W_b) dt \right\|_q & \leq 4\beta_q \sqrt{T-b} \|W_b - W_b^{(a,b]}\|_q \\ & \leq 8\beta_q \sqrt{T-b} \|W_b - W_a\|_q \\ & \leq 8\beta_q^2 \sqrt{T-b} \sqrt{b-a}. \end{aligned}$$

Summarizing gives

$$\|\xi - \xi^{(a,b]}\|_p \leq \beta_p \left[2\sqrt{b-a} + (8\beta_q^2 \sqrt{T-b} \sqrt{b-a})^{\frac{1}{2}} \right].$$

(b) Let us turn to the lower bound, where we assume $p \in (1, \infty)$. We obtain

$$\begin{aligned} & \|\xi - \xi^{(a,b]}\|_p \\ & \geq -2 \left\| \int_{(a,b]} \chi_{\{W_t > \alpha\}} dW_t \right\|_p + \left\| \int_{(b,T]} [\chi_{\{W_t > \alpha\}} - \chi_{\{W_t^{(a,b]} > \alpha\}}] dW_t \right\|_p \\ & \geq -2\beta_p \sqrt{b-a} + \frac{1}{\beta_p} \left\| \int_b^T \chi_{I_\alpha(W_b, W_b^{(a,b]})}(W_t - W_b) dt \right\|_q^{\frac{1}{2}}. \end{aligned}$$

Let $a = 0$ and observe that on $\{W_b \leq -\sqrt{b}, W'_b \geq \sqrt{b}\}$ one has that

$$I_\alpha(W_b, W_b^{(0,b]}) = (\alpha - W'_b, \alpha - W_b] \supseteq (\alpha - \sqrt{b}, \alpha + \sqrt{b}).$$

Therefore, for $b \in (0, T/2)$,

$$\begin{aligned} & \|\xi - \xi^{(0,b]}\|_p \\ & \geq -2\beta_p \sqrt{b} + \frac{1}{\beta_p} \bar{\mathbb{P}}(W_b \leq -\sqrt{b}, W'_b \geq \sqrt{b})^{\frac{1}{2q}} \left\| \int_b^T \chi_{(\alpha - \sqrt{b}, \alpha + \sqrt{b})}(W_t - W_b) dt \right\|_q^{\frac{1}{2}} \\ & = -2\beta_p \sqrt{b} + \frac{1}{\beta_p} \bar{\mathbb{P}}(W_1 \leq -1, W'_1 \geq 1)^{\frac{1}{2q}} \left\| \int_b^T \chi_{(\alpha - \sqrt{b}, \alpha + \sqrt{b})}(W_t - W_b) dt \right\|_q^{\frac{1}{2}} \end{aligned}$$

$$\begin{aligned}
&\geq -2\beta_p\sqrt{b} + \frac{1}{\beta_p}\overline{\mathbb{P}}(W_1 \leq -1, W'_1 \geq 1)^{\frac{1}{2q}} \left\| \int_b^{\frac{T}{2}+b} \chi_{(\alpha-\sqrt{b}, \alpha+\sqrt{b})}(W_t - W_b) dt \right\|_q^{\frac{1}{2}} \\
&= -2\beta_p\sqrt{b} + \frac{1}{\beta_p}\overline{\mathbb{P}}(W_1 \leq -1, W'_1 \geq 1)^{\frac{1}{2q}} \left\| \int_0^{\frac{T}{2}} \chi_{(\alpha-\sqrt{b}, \alpha+\sqrt{b})}(W_t) dt \right\|_q^{\frac{1}{2}}.
\end{aligned}$$

For the local time of the Brownian motion one has (see [69, Corollary VI.1.9])

$$L_t^\alpha = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t \chi_{(\alpha-\varepsilon, \alpha+\varepsilon)}(W_s) ds \text{ a.s.}$$

Therefore, by Fatou's Lemma,

$$\liminf_{b \downarrow 0} \frac{1}{\sqrt[4]{b}} \left\| \int_0^{\frac{T}{2}} \chi_{(\alpha-\sqrt{b}, \alpha+\sqrt{b})}(W_t) dt \right\|_q^{\frac{1}{2}} \geq \sqrt{2\|L_T^\alpha\|_q} > 0.$$

□

Because the local time L_t^α can be expressed by Tanaka's formula by

$$\frac{1}{2}L_T^\alpha = (W_T - \alpha)^+ - (W_0 - \alpha)^+ - N_T^\alpha,$$

see [69, Theorem VI.1.2], and because $(W_T - \alpha)^+ \in \mathbb{B}_p^{\Phi_2}$ for all $p \in (0, \infty)$ we immediately get the following corollary:

COROLLARY 4.27. *For all $\alpha \in \mathbb{R}$ and $p \in (1, \infty)$ one has that*

$$L_T^\alpha \in \mathbb{B}_p^{\Phi_4} \setminus \left[\bigcup_{r \in [2, 4)} \mathbb{B}_p^{\Phi_r} \right].$$

CHAPTER 5

BMO-Martingales

The theory of BMO-martingales has become an important tool in the investigation of BSDEs. For an account on this topic the reader is referred, for example, to [29, p. 298] and [19, p. 2922]. In particular, there are two key ingredients that we will use as well: Fefferman's inequality and their generalizations, and the notion of reverse Hölder inequalities. In addition to these two ingredients, we exploit the concept of sliceable BMO-martingales which can be seen as a natural enhancement for the previous techniques. Sliceable BMO-martingales were used by Emery [31, 32] and Schachermayer [70], and in the context of backward stochastic differential equations by Delbaen and Tang [29] and Frei [33].

Throughout this chapter we assume a stochastic basis $(A, \mathcal{A}, \mathbb{Q}, (\mathcal{A}_t)_{t \in [0, T]})$, $T > 0$, where $(A, \mathcal{A}, \mathbb{Q})$ is complete, $(\mathcal{A}_t)_{t \in [0, T]}$ is right-continuous, \mathcal{A}_0 contains all null-sets, and $\mathcal{A} = \mathcal{A}_T$. To be in accordance with [50], we assume that all local martingales are continuous. As we work on a closed time-interval we have to explain our understanding of a local martingale: we require that the localizing sequence of stopping times $0 \leq \tau_1 \leq \tau_2 \leq \dots \leq T$ satisfies $\lim_n \mathbb{P}(\tau_n = T) = 1$. So we extend the filtration by \mathcal{A}_T to (T, ∞) , i.e. $\mathcal{A}_t := \mathcal{A}_T$ for $t \in (T, \infty)$, and extend all local martingales $(N_t)_{t \in [0, T]}$ (in our setting) by N_T to (T, ∞) . This yields the standard notion of a local martingale.

5.1. BMO-martingales

First we recall the notion of a BMO-martingale.

DEFINITION 5.1. A martingale $M = (M_t)_{t \in [0, T]}$ is of *bounded mean oscillation* (we write $M \in \text{BMO}$) provided that $M_0 \equiv 0$ and there is constant $c > 0$ such that for all stopping times $\tau : A \rightarrow [0, T]$ one has that

$$\mathbb{E}(|M_T - M_\tau|^2 | \mathcal{A}_\tau) \leq c^2 \text{ a.s..}$$

We let $\|M\|_{\text{BMO}_2} := \inf c$ where the infimum is taken over all $c > 0$ as above.

Next we introduce the *sliceable numbers*. Without being defined explicitly, these numbers have their origin in an article of Schachermayer [70] and will be used via Theorem 5.9 below in our article. Before giving the definition let us recall the notation

$${}^\sigma M^\tau := (M_{\tau \wedge t} - M_{\sigma \wedge t})_{t \in [0, T]}.$$

DEFINITION 5.2. For a BMO-martingale $M = (M_t)_{t \in [0, T]}$ and $N \geq 1$ we let

$$\text{sl}_N(M) := \inf \varepsilon,$$

where the infimum is taken over all $\varepsilon > 0$ such that there are stopping times $0 = \tau_0 \leq \tau_1 \leq \dots \leq \tau_N = T$ with

$$\sup_{k=1, \dots, N} \|\tau_{k-1} M^{\tau_k}\|_{\text{BMO}_2} \leq \varepsilon.$$

Moreover, we let

$$\text{sl}_\infty(M) := \lim_N \text{sl}_N(M).$$

We call $\text{sl}_N(M)$ the N -sliceable number of M . The BMO-martingale M is called sliceable provided that $\text{sl}_\infty(M) = 0$.

Before we summarise some simple properties of the sliceable numbers we need the following lemma:

LEMMA 5.3. Let $0 \leq \sigma \leq \tau \leq T$ be stopping times and $0 = \tau_0 \leq \tau_1 \leq \dots \leq \tau_N = T$ be a net of stopping times such that for all $\omega \in A$ there is a $k \in \{1, \dots, N\}$ such that

$$(\sigma(\omega), \tau(\omega)] \subseteq (\tau_{k-1}(\omega), \tau_k(\omega)].$$

Then, for a BMO-martingale N , one has that

$$\|\sigma N^\tau\|_{\text{BMO}_2} \leq \sup_{k=1, \dots, N} \|\tau_{k-1} N^{\tau_k}\|_{\text{BMO}_2}.$$

PROOF. Let $\rho : A \rightarrow [0, T]$ be a stopping time. Then

$$\begin{aligned} \mathbb{E}(|^\sigma N_T^\tau - {}^\sigma N_\rho^\tau|^2 | \mathcal{A}_\rho) &= \mathbb{E}(|N_{\tau \vee \rho} - N_{\sigma \vee \rho}|^2 | \mathcal{A}_\rho) \\ &= \mathbb{E}(\mathbb{E}(|N_{\tau \vee \rho} - N_{\sigma \vee \rho}|^2 | \mathcal{A}_{\sigma \vee \rho}) | \mathcal{A}_\rho). \end{aligned}$$

Now we observe that $(\bar{\sigma}, \bar{\tau})$ with $\bar{\sigma} := \sigma \vee \rho$ and $\bar{\tau} := \tau \vee \rho$ shares the same property as (σ, τ) . We let $A_{N+1} := \{\bar{\sigma} = T\}$, and for $k = 1, \dots, N$,

$$A_k := \{\bar{\sigma} \in [\tau_{k-1}, \tau_k)\}.$$

This gives a partition $A = \bigcup_{k=1}^{N+1} A_k$ with $A_k \in \mathcal{A}_{\bar{\sigma}}$ and we have that

$$\begin{aligned} \mathbb{E}(|N_{\bar{\tau}} - N_{\bar{\sigma}}|^2 | \mathcal{A}_{\bar{\sigma}}) &= \sum_{k=1}^N \mathbb{E}(\chi_{A_k} |N_{\bar{\tau}} - N_{\bar{\sigma}}|^2 | \mathcal{A}_{\bar{\sigma}}) \\ &= \sum_{k=1}^N \mathbb{E}(\chi_{A_k} |N_{\bar{\tau} \wedge \tau_k} - N_{\bar{\sigma} \vee \tau_{k-1}}|^2 | \mathcal{A}_{\bar{\sigma}}) \\ &= \sum_{k=1}^N \mathbb{E}(\chi_{A_k} \mathbb{E}(|N_{\bar{\tau} \wedge \tau_k} - N_{\bar{\sigma} \vee \tau_{k-1}}|^2 | \mathcal{A}_{\bar{\sigma} \vee \tau_{k-1}}) | \mathcal{A}_{\bar{\sigma}}) \\ &\leq \sup_{k=1, \dots, N} \|\tau_{k-1} N^{\tau_k}\|_{\text{BMO}_2}^2. \end{aligned}$$

□

To formulate the next result we recall the space \mathbb{H}_∞ :

DEFINITION 5.4. We let \mathbb{H}_∞ be the space of all (continuous) martingales $N = (N_t)_{t \in [0, T]}$ such that $N_0 \equiv 0$ and

$$\|N\|_{\mathbb{H}_\infty} := \text{esssup}_{\omega \in A} [N, N]_T(\omega) < \infty.$$

It follows directly from the definition that $\mathbb{H}_\infty \subseteq \text{BMO}$.

LEMMA 5.5. *For BMO-martingales M , M_1 , and M_2 one has the following:*

- (1) $\text{sl}_1(M) = \|M\|_{\text{BMO}_2}$.
- (2) $\text{sl}_1(M) \geq \text{sl}_2(M) \geq \dots \geq 0$.
- (3) $\text{sl}_{N_1+N_2-1}(M_1 + M_2) \leq \text{sl}_{N_1}(M_1) + \text{sl}_{N_2}(M_2)$.
- (4) $\text{sl}_\infty(M) = d_{\text{BMO}_2}(M, \mathbb{H}_\infty)$, where

$$d_{\text{BMO}_2}(M, \mathbb{H}_\infty) := \inf\{\|M - N\|_{\text{BMO}_2} : N \in \mathbb{H}_\infty\}.$$

PROOF. (1) and (2) are obvious. To prove (3), we assume $\eta > 0$ and find nets $0 = \tau_0^i \leq \dots \leq \tau_{N_i}^i = T$ such that

$$\sup_{k=1, \dots, N_i} \left\| \tau_{k-1}^i M_i^{\tau_k^i} \right\|_{\text{BMO}_2} \leq \text{sl}_{N_i}(M_i) + \eta.$$

Now we let $(\sigma_k)_{k=0}^{N_1+N_2-1}$ be the union of $(\tau_k^1)_{k=0}^{N_1}$ and $(\tau_k^2)_{k=0}^{N_2}$ and define the new net $(\tau_k)_{k=0}^{N_1+N_2-1}$ to be the order statistics of $(\sigma_k)_{k=0}^{N_1+N_2-1}$, i.e.

$$\begin{aligned} \tau_0 &:= \min_k \sigma_k = 0, \\ \tau_{N_1+N_2-1} &:= \max_k \sigma_k = T, \\ \tau_k &:= \min_{\substack{I \subseteq \{1, \dots, N_1+N_2-2\} \\ \text{card}(I) = k}} \max_{l \in I} \sigma_l. \end{aligned}$$

With this definition and Lemma 5.3 we get for $k = 1, \dots, N_1 + N_2 - 1$ that

$$\begin{aligned} & \left\| \tau_{k-1}^{N_1+N_2-1} (M_1 + M_2)^{\tau_k} \right\|_{\text{BMO}_2} \\ & \leq \left\| \tau_{k-1}^{N_1+N_2-1} M_1^{\tau_k} \right\|_{\text{BMO}_2} + \left\| \tau_{k-1}^{N_1+N_2-1} M_2^{\tau_k} \right\|_{\text{BMO}_2} \\ & \leq \sup_{k_1=1, \dots, N_1} \left\| \tau_{k_1-1}^1 M_1^{\tau_{k_1}^1} \right\|_{\text{BMO}_2} + \sup_{k_2=1, \dots, N_2} \left\| \tau_{k_2-1}^2 M_2^{\tau_{k_2}^2} \right\|_{\text{BMO}_2} \\ & \leq \text{sl}_{N_1}(M_1) + \text{sl}_{N_2}(M_2) + 2\eta. \end{aligned}$$

By $\eta \downarrow 0$ the assertion follows.

(4) This part follows from [70] where we have to observe that our setting of a bounded time interval $[0, T]$ does not make a difference compared to $[0, \infty)$ from [70]. \square

The next example will be used later:

EXAMPLE 5.6. Assume that

$$\langle M \rangle_t = \int_0^t c_s^2 ds, \quad t \in [0, T], \quad \text{a.s.}$$

for some predictable process $c = (c_t)_{t \in [0, T]}$ and that there is a $\delta > 0$ and some $\kappa \in [0, \infty)$ such that

$$\left[\mathbb{E} \left(\int_\tau^T |c_s|^{2+\delta} ds \middle| \mathcal{A}_\tau \right) \right]^{\frac{1}{2+\delta}} \leq \kappa \text{ a.s.}$$

for all stopping times $\tau : A \rightarrow [0, T]$. Then, for $\alpha := \frac{1}{2} - \frac{1}{2+\delta} > 0$, and $N \geq 1$,

$$\text{sl}_N(M) \leq \kappa \left(\frac{T}{N} \right)^\alpha.$$

PROOF. For $0 \leq a < b \leq T$ we simply get a.s. that

$$\begin{aligned}
\left[\mathbb{E} \left(\int_{\tau}^T \chi_{(a,b]}(s) |c_s|^2 ds \middle| \mathcal{A}_{\tau} \right) \right]^{\frac{1}{2}} &= \left[\mathbb{E} \left(\int_{\tau \vee a}^{\tau \vee b} |c_s|^2 ds \middle| \mathcal{A}_{\tau} \right) \right]^{\frac{1}{2}} \\
&\leq \left[\mathbb{E} \left(\int_{\tau \vee a}^{\tau \vee b} |c_s|^{2+\delta} ds \middle| \mathcal{A}_{\tau} \right) \right]^{\frac{1}{2+\delta}} (b-a)^{\alpha} \\
&\leq \left[\mathbb{E} \left(\int_{\tau}^T |c_s|^{2+\delta} ds \middle| \mathcal{A}_{\tau} \right) \right]^{\frac{1}{2+\delta}} (b-a)^{\alpha} \\
&\leq \kappa (b-a)^{\alpha}.
\end{aligned}$$

Choosing an equidistant partition of $[0, T]$ consisting of N intervals concludes the proof. \square

5.2. Reverse Hölder inequalities

The probabilistic Muckenhoupt weights provide a natural way to verify various martingale inequalities after a change of measure, see exemplary [48, 12, 50]. This change of measure will appear in our setting in terms of a Girsanov transformation that removes (see Section 6.6) a sub-quadratic or quadratic drift term in Z that originates from the generator of our BSDE.

DEFINITION 5.7. Assume a martingale $M = (M_t)_{t \in [0, T]}$ with $M_0 \equiv 0$ such that $\mathcal{E}(M)$ with

$$\mathcal{E}(M)_t = e^{M_t - \frac{1}{2} \langle M \rangle_t}$$

for $t \in [0, T]$ is a martingale as well. For $\beta \in (1, \infty)$ we let $\mathcal{E}(M) \in \mathcal{RH}_{\beta}$ provided that there is a constant $c > 0$ such that for all stopping times $\tau : A \rightarrow [0, T]$ one has that

$$\mathbb{E}(|\mathcal{E}(M)_T|^{\beta} | \mathcal{A}_{\tau})^{\frac{1}{\beta}} \leq c \mathcal{E}(M)_{\tau} \text{ a.s.}$$

The smallest possible $c \geq 0$ is denoted by $\mathcal{RH}_{\beta}(\mathcal{E}(M))$.

It is known [50, Theorem 2.3] that $\mathcal{E}(M)$ is a martingale for $M \in \text{BMO}$. Moreover, we have the following result:

PROPOSITION 5.8 ([50, Theorems 2.4 and 3.4]). *Let M be a martingale with $M_0 \equiv 0$ such that $\mathcal{E}(M)$ is a martingale. Then $M \in \text{BMO}$ if and only if $\mathcal{E}(M) \in \bigcup_{\beta \in (1, \infty)} \mathcal{RH}_{\beta}$.*

Later in our application we need to know whether a certain martingale M generates a Doléan-Dade exponential that satisfies a reverse Hölder inequality. Here the BMO_2 -distance to L_{∞} would be a natural candidate for the extreme case that the reverse Hölder inequality is satisfied for all parameters $\beta \in (1, \infty)$, as Kazamaki [50, Theorem 3.8] provides the characterization $M \in \overline{L_{\infty}}^{[\text{BMO}, \|\cdot\|_{\text{BMO}_2}]}$ for this case. On the other hand, Grandits [41] has shown that a positive BMO_2 -distance to L_{∞} does not provide a reasonable estimate for the critical value of β such that one has a reverse Hölder inequality (see also the *Note added in Proof* of [70]). This is our reason to use the concept *sliceable* (which describes the BMO_2 -distance to \mathbb{H}_{∞} due to the result of Schachermayer [70]) because the following observation yields

explicit estimates for the critical exponent β and the corresponding multiplicative constants in the reverse Hölder inequalities:

THEOREM 5.9. *Let $\Phi : (1, \infty) \rightarrow (0, \infty)$ be a non-increasing function and let*

$$\Psi : \left\{ (\gamma, \beta) \in [0, \infty) \times (1, \infty) : 0 \leq \gamma < \Phi(\beta) < \infty \right\} \rightarrow [0, \infty)$$

be right-continuous in its first argument and such that

$$\Psi(\gamma_1, \beta) \leq \Psi(\gamma_2, \beta) \quad \text{for } 0 \leq \gamma_1 \leq \gamma_2 < \Phi(\beta),$$

with the property that

$$\|M\|_{\text{BMO}_2} < \Phi(\beta) \quad \text{implies} \quad \mathcal{RH}_\beta(\mathcal{E}(M)) \leq \Psi(\|M\|_{\text{BMO}_2}, \beta).$$

Then, for $\text{sl}_N(M) < \Phi(\beta)$ we have that $\mathcal{RH}_\beta(\mathcal{E}(M)) \leq [\Psi(\text{sl}_N(M), \beta)]^N$.

PROOF. The proof is based on a simple recursion argument that uses the concept of a sliceable BMO-martingale. For $\text{sl}_N(M) < \Phi(\beta)$ we choose $0 = \tau_0 \leq \dots \leq \tau_N = T$ such that

$$\|\tau_{k-1} M^{\tau_k}\|_{\text{BMO}_2} < \text{sl}_N(M) + \eta < \Phi(\beta)$$

for some $\eta > 0$ and all $k = 1, \dots, N$. Therefore,

$$\mathcal{RH}_\beta(\mathcal{E}(\tau_{k-1} M^{\tau_k})) \leq \Psi(\|\tau_{k-1} M^{\tau_k}\|_{\text{BMO}_2}, \beta) \leq \Psi(\text{sl}_N(M) + \eta, \beta).$$

Letting $\tau : A \rightarrow [0, T]$ be a stopping time and $\sigma_k := \tau_k \vee \tau$ gives that

$$\begin{aligned} & \mathbb{E}_{\mathcal{A}_\tau} \left(e^{\beta(M_T - \frac{1}{2}\langle M \rangle_T)} \right) \\ &= \left(e^{\beta(M_\tau - \frac{1}{2}\langle M \rangle_\tau)} \right) \mathbb{E}_{\mathcal{A}_\tau} \left(e^{\beta([M_T - M_\tau] - \frac{1}{2}[\langle M \rangle_T - \langle M \rangle_\tau])} \right) \\ &= \left(e^{\beta(M_\tau - \frac{1}{2}\langle M \rangle_\tau)} \right) \mathbb{E}_{\mathcal{A}_\tau} \left(\prod_{k=1}^N e^{\beta([M_{\sigma_k} - M_{\sigma_{k-1}}] - \frac{1}{2}[\langle M \rangle_{\sigma_k} - \langle M \rangle_{\sigma_{k-1}}])} \right). \end{aligned}$$

Next we observe that

$$(17) \quad \mathbb{E}_{\mathcal{A}_{\sigma_{k-1}}} \left(e^{\beta([M_{\sigma_k} - M_{\sigma_{k-1}}] - \frac{1}{2}[\langle M \rangle_{\sigma_k} - \langle M \rangle_{\sigma_{k-1}}])} \right) \leq [\Psi(\text{sl}_N(M) + \eta, \beta)]^\beta$$

for $k = 1, \dots, N$ which follows from

$$\|\sigma_{k-1} M^{\sigma_k}\|_{\text{BMO}_2} = \|\tau \vee \tau_{k-1} M^{\tau \vee \tau_k}\|_{\text{BMO}_2} \leq \sup_{l=1, \dots, N} \|\tau_{l-1} M^{\tau_l}\|_{\text{BMO}_2} < \text{sl}_N(M) + \eta,$$

where we use Lemma 5.3. Applying (17) inductively backwards beginning with $k = N$ and using the projection property of the conditional expectation gives that

$$\mathcal{RH}_\beta(\mathcal{E}(M))^\beta \leq [\Psi(\text{sl}_N(M) + \eta, \beta)]^{\beta N}.$$

We conclude by $\eta \downarrow 0$. □

According to [50, Proof of Theorem 3.1] possible choices of (Φ, Ψ) are

$$(18) \quad \Phi(\beta) := \left(1 + \frac{1}{\beta^2} \log \left(1 + \frac{1}{2\beta - 2} \right) \right)^{\frac{1}{2}} - 1,$$

$$(19) \quad \Psi(\gamma, \beta) := \left(\frac{2}{1 - \frac{2\beta - 2}{2\beta - 1} e^{\beta^2[\gamma^2 + 2\gamma]}} \right)^{\frac{1}{\beta}},$$

where Φ is decreasing with $\lim_{\beta \rightarrow \infty} \Phi(\beta) = 0$ and $\lim_{\beta \rightarrow 1} \Phi(\beta) = \infty$.

5.3. Fefferman's inequality and $\text{BMO}(S_{2\theta})$ spaces

In this section we slightly change the point of view: Instead of considering martingales we think in terms of the quadratic variation which is more convenient in the sequel for us. The BMO-spaces, related to backward stochastic differential equations with generators satisfying condition (B3) of Section 6.1 below, are defined as follows:

DEFINITION 5.10. For $\theta \in (0, \infty)$ and an \mathbb{R} -valued progressively measurable process $Z = (Z_t)_{t \in [0, T]}$ with $\mathbb{E} \int_0^T |Z_s|^{2\theta} ds < \infty$ we let $Z \in \text{BMO}(S_{2\theta})$ provided that

$$\|Z\|_{\text{BMO}(S_{2\theta})} := \sup_{t \in [0, T]} \left\| \mathbb{E} \left(\int_t^T |Z_s|^{2\theta} ds \middle| \mathcal{A}_t \right) \right\|_{\infty}^{\frac{1}{2\theta}} < \infty.$$

The notation $S_{2\theta}$ is chosen to indicate that $\text{BMO}(S_{2\theta})$ deals with a modified square function. For $\theta \in (1, \infty)$ we obtain a condition that is stronger than the classical BMO-condition $\|Z\|_{\text{BMO}(S_2)}$, whereas for $\theta \in (0, 1)$ the condition gets weaker. If we define

$$A_t := \int_0^t |Z_s|^{2\theta} ds,$$

then $Z \in \text{BMO}(S_{2\theta})$ if and only if

$$\sup_{\tau} \|\mathbb{E}(A_T - A_{\tau} | \mathcal{F}_{\tau})\|_{\infty} < \infty$$

with the supremum taken over all stopping times $\tau : A \rightarrow [0, T]$. This opens the path to apply known results about BMO-spaces to the $\text{BMO}(S_{2\theta})$ -spaces. Therefore, by the John-Nirenberg Theorem we get that $Z \in \text{BMO}(S_{2\theta})$ implies that

$$(20) \quad \int_0^T |Z_s|^{2\theta} ds \in L_{\text{exp}},$$

where the Orlicz space L_{exp} is given by

$$\|F\|_{L_{\text{exp}}} := \inf \left\{ \lambda > 0 : \mathbb{E} e^{\frac{|F|}{\lambda}} \leq 2 \right\}$$

for a random variable F taking values in \mathbb{R} , see [71, 34, 50] and [37, Corollary 1].

For the next example the notion of a Banach function space is convenient:

DEFINITION 5.11. A map $\rho : \mathcal{L}_0^+(A, \mathcal{A}, \mathbb{Q}) \rightarrow [0, \infty]$ defined on the non-negative random variables of $\mathcal{L}_0(A, \mathcal{A}, \mathbb{Q})$ is a Banach function norm provided that the following conditions are satisfied:

- (1) $\rho(X) = 0$ if and only if $X = 0$ a.s.
- (2) $\rho(X + Y) \leq \rho(X) + \rho(Y)$.
- (3) $\rho(\alpha X) = \alpha \rho(X)$ for $\alpha \geq 0$.
- (4) $0 \leq X \leq Y$ a.s. implies $\rho(X) \leq \rho(Y)$.
- (5) $0 \leq X_n \uparrow X$ a.s. implies $\rho(X_n) \uparrow \rho(X)$.
- (6) $\rho(1) < \infty$.
- (7) There is a $c > 0$ such that $\|X\|_1 \leq c\rho(X)$ for all $X \in \mathcal{L}_0^+(A, \mathcal{A}, \mathbb{Q})$.

The function ρ is extended to $\|\cdot\|_\rho : L_0(A, \mathcal{A}, \mathbb{Q}) \rightarrow [0, \infty]$ by $\|X\|_\rho := \rho(|X|)$ and we let

$$E_\rho := \{X \in L_0(A, \mathcal{A}, \mathbb{Q}) : \|X\|_\rho < \infty\}.$$

The spaces $[E_\rho, \|\cdot\|_\rho]$ are Banach spaces having the Fatou property, see [8, Theorem 1.1.7].

EXAMPLE 5.12. Let $T = 1$ and assume that $\rho : L_0(A, \mathcal{A}, \mathbb{Q}) \rightarrow [0, \infty]$ is a Banach function norm such that for all $t \in (0, 1]$ one has that

$$\sup\{\|X\|_\infty : \|X\|_\rho \leq 1, X \in L_0(A, \mathcal{A}_t, \mathbb{Q})\} = \infty.$$

Then for all $0 < \theta < \eta \leq 1$ there is a progressively measurable process $Z = (Z_t)_{t \in [0, T]}$ such that

- (1) $\int_0^T |Z_t|^{2\eta} dt \in E_\rho$,
- (2) $Z \in \text{BMO}(S_{2\theta}) \setminus \text{BMO}(S_{2\eta})$.

PROOF. Let $t_n := 1 - \frac{1}{2^n}$ for $n \geq 0$, take

$$0 < \varepsilon < \frac{1}{2\theta} - \frac{1}{2\eta},$$

and choose, for $n \geq 1$ random variables $v_n : A \rightarrow \mathbb{R}$ that are \mathcal{A}_{t_n} -measurable and satisfy

$$\|v_n\|_\infty = 2^{(n+1)[\frac{1}{2\eta} + \varepsilon]} \quad \text{but} \quad \| |v_n|^{2\eta} \|_{E_\rho} \leq 1.$$

Define the stochastic process $Z = (Z_t)_{t \in [0, 1]}$ by

$$Z_t := \sum_{n=2}^{\infty} \chi_{(t_{n-1}, t_n]}(t) v_{n-1}.$$

Then we get the following three estimates:

- (1) For $n \geq 2$ we have

$$\|Z\|_{\text{BMO}(S_{2\eta})} \geq \|v_{n-1}\|_\infty (t_n - t_{n-1})^{\frac{1}{2\eta}} = 2^{n[\frac{1}{2\eta} + \varepsilon]} 2^{-\frac{n}{2\eta}} \rightarrow \infty$$

as $n \rightarrow \infty$, so that $\|Z\|_{\text{BMO}(S_{2\eta})} = \infty$.

- (2) We have that

$$\begin{aligned} \|Z\|_{\text{BMO}(S_{2\theta})}^{2\theta} &\leq \sum_{n=2}^{\infty} \|v_{n-1}\|_\infty^{2\theta} (t_n - t_{n-1}) \\ &= \sum_{n=2}^{\infty} 2^{n[\frac{\theta}{\eta} + 2\varepsilon\theta - 1]} < \infty. \end{aligned}$$

- (3) On the other side, we have that

$$\begin{aligned} \left\| \int_0^T |Z_t|^{2\eta} dt \right\|_{E_\rho} &\leq \sum_{n=2}^{\infty} \| |v_{n-1}|^{2\eta} \|_{E_\rho} (t_n - t_{n-1}) \\ &\leq \sum_{n=2}^{\infty} 2^{-n} < \infty. \end{aligned}$$

□

In the following we give a version of the generalized Fefferman's inequality that can be found in [29, Lemma 1.6], see also [4, Theorem 1.1]. Our contribution in Theorem 5.20 below consists in improving the asymptotic behavior of the constant from p to \sqrt{p} in Corollary 5.21 and that the left-hand side in (21) is stronger than the left-hand side in (22).

We start with the definition of the $\mathbb{H}_p(S_2)$ -spaces and continue by some elementary lemmas.

DEFINITION 5.13. For $p \in (0, \infty]$ we define $\mathbb{H}_p(S_2)$ to be the space of all progressively measurable \mathbb{R} -valued process $A = (A_t)_{t \in [0, T]}$ such that

$$\|A\|_{\mathbb{H}_p(S_2)} := \left\| \left(\int_0^T |A_s|^2 ds \right)^{\frac{1}{2}} \right\|_p < \infty.$$

LEMMA 5.14. Let μ be a finite measure on $\mathcal{B}([0, T])$ with $\mu([0, T]) > 0$, $\theta \in (0, 1)$, and let

$$t_0 := \inf\{t \in [0, T] : \mu([0, t]) > 0\}.$$

Then one has that

$$\int_{[t_0, T]} \mu([0, t])^{\theta-1} d\mu(t) \leq \frac{1}{\theta} \mu([0, T])^\theta.$$

The proof is standard and we leave it to the reader.

LEMMA 5.15. Let $p \in (1, \infty)$, ν be a finite measure on $\mathcal{B}([0, T])$, and $f : [0, T] \rightarrow [0, \infty)$ be non-decreasing and right-hand side continuous. Then

$$\left| \int_{[0, T]} f(s) d\nu(s) \right|^p \leq p \int_{[0, T]} \left| \int_{[0, t]} f(s) d\nu(s) \right|^{p-1} f(t) d\nu(t).$$

PROOF. For $n \geq 1$ take the equi-spaced grid

$$0 = t_0^n < t_1^n < \dots < t_{2^n}^n = T.$$

By dominated convergence it is enough to show that

$$\begin{aligned} & \left| f(0)\nu(\{0\}) + \sum_{i=1}^{2^n} f(t_i^n)\nu((t_{i-1}^n, t_i^n]) \right|^p \\ & \leq p |f(0)\nu(\{0\})|^{p-1} f(0)\nu(\{0\}) + \\ & \quad p \sum_{i=1}^{2^n} \left(f(0)\nu(\{0\}) + \sum_{j=1}^i f(t_j^n)\nu((t_{j-1}^n, t_j^n]) \right)^{p-1} f(t_i^n)\nu((t_{i-1}^n, t_i^n]). \end{aligned}$$

Setting $a_0 := f(0)\nu(\{0\})$ and $a_i := f(t_i^n)\nu((t_{i-1}^n, t_i^n])$ for $i = 1, \dots, 2^n$, this reads as

$$\left| \sum_{i=0}^{2^n} a_i \right|^p \leq p \sum_{i=0}^{2^n} \left(\sum_{j=0}^i a_j \right)^{p-1} a_i$$

which follows by writing the left-hand side as telescoping sum and applying the mean-value theorem from calculus. \square

REMARK 5.16. In Lemmas 5.14 and 5.15 the factors $1/\theta$ and p are sharp, but one does not have equalities in general (one can check the cases where μ and ν are either the Lebesgue measure or the Dirac measure at (say) T , and $f \equiv 1$).

DEFINITION 5.17. We call a map

$$\nu : A \times \mathcal{B}([0, T]) \rightarrow [0, \infty)$$

adapted random measure provided that

- (1) the map $\nu(\omega, \cdot) : \mathcal{B}([0, T]) \rightarrow [0, \infty)$ is a measure for all $\omega \in A$,
- (2) the map $\nu(\cdot, [0, t]) : A \rightarrow \mathbb{R}$ is \mathcal{A}_t -measurable for all $t \in [0, T]$.

Moreover, we let

$$\|\nu\|_{BMO} := \sup_{t \in [0, T]} \|\mathbb{E}(\nu([t, T]) | \mathcal{A}_t)\|_{\infty}.$$

Given any non-negative, non-decreasing, left-hand side continuous, and adapted process $(f(s))_{s \in [0, T]}$, the process $(\int_{[0, t]} f(s) d\nu(s))_{t \in [0, T]}$ is well defined, non-decreasing, right-hand side continuous, and adapted.

LEMMA 5.18. *Let ν be an adapted random measure, $(f(t))_{t \in [0, T]}$ be non-decreasing, adapted, non-negative, and left-hand side continuous. Then, one has that*

$$\mathbb{E} \int_{[0, T]} f(s) d\nu(s) \leq \mathbb{E} f(T) \|\nu\|_{BMO}.$$

PROOF. We can assume that $\mathbb{E} f(T) \|\nu\|_{BMO} < \infty$, otherwise there is nothing to prove. Assuming the equi-spaced net

$$0 = t_0^n < \dots < t_{2^n}^n = T,$$

it is sufficient to show that

$$\mathbb{E} \sum_{i=0}^{2^n-1} f(t_i^n) \nu([t_i^n, t_{i+1}^n)) + \mathbb{E} f(T) \nu(\{T\}) \leq \mathbb{E} f(T) \sup_{j=0, \dots, 2^n} \|\mathbb{E}(\nu([t_j^n, T]) | \mathcal{A}_{t_j^n})\|_{\infty}.$$

Letting $q_i^n := \nu([t_i^n, t_{i+1}^n))$ for $i = 0, \dots, 2^n - 1$, $q_{2^n}^n := \nu(\{T\})$, and $a_0^n + \dots + a_i^n = f(t_i^n)$, we get that

$$\begin{aligned} & \mathbb{E} \left[\sum_{i=0}^{2^n-1} f(t_i^n) \nu([t_i^n, t_{i+1}^n)) + f(T) \nu(\{T\}) \right] \\ &= \mathbb{E} \left[\sum_{0 \leq j \leq i \leq 2^n} a_j^n q_i^n \right] \\ &= \sum_{j=0}^{2^n} \mathbb{E} \left[a_j^n \mathbb{E}(q_j^n + \dots + q_{2^n}^n | \mathcal{A}_{t_j^n}) \right] \\ &\leq \mathbb{E} f(T) \sup_{j=0, \dots, 2^n} \|\mathbb{E}(\nu([t_j^n, T]) | \mathcal{A}_{t_j^n})\|_{\infty}. \end{aligned}$$

□

LEMMA 5.19. *Let μ and ν be adapted random measures such that $(\mu(\cdot, [0, t]))_{t \in [0, T]}$ and $(\nu(\cdot, [0, t]))_{t \in [0, T]}$ are continuous processes. Let $\eta \in (0, 1)$, $p \in (1, \infty)$, and assume that*

$$\mathbb{E} \left| \int_{[0, T]} \mu([0, t])^\eta d\nu(t) \right|^p < \infty.$$

Then we have that

$$\left\| \int_{[0,T]} \mu([0,t])^\eta d\nu(t) \right\|_p \leq p \|\mu([0,T])^\eta\|_p \|\nu\|_{\text{BMO}}.$$

PROOF. For $p \in (1, \infty)$ we use Lemma 5.15 and Lemma 5.18 to get that

$$\begin{aligned} & \mathbb{E} \left| \int_{[0,T]} \mu([0,t])^\eta d\nu(t) \right|^p \\ & \leq p \mathbb{E} \int_{[0,T]} \left| \int_{[0,t]} \mu([0,s])^\eta d\nu(s) \right|^{p-1} \mu([0,t])^\eta d\nu(t) \\ & \leq p \mathbb{E} \left[\left| \int_{[0,T]} \mu([0,s])^\eta d\nu(s) \right|^{p-1} \mu([0,T])^\eta \right] \|\nu\|_{\text{BMO}} \\ & \leq p \left[\mathbb{E} \left| \int_{[0,T]} \mu([0,t])^\eta d\nu(t) \right|^p \right]^{\frac{p-1}{p}} \|\mu([0,T])^\eta\|_p \|\nu\|_{\text{BMO}}. \end{aligned}$$

Dividing by $\left[\mathbb{E} \left| \int_{[0,T]} \mu([0,t])^\eta d\nu(t) \right|^p \right]^{\frac{p-1}{p}}$ in the case this expression is positive (otherwise there is nothing to prove), gives the desired inequality. \square

THEOREM 5.20. Let μ, ν be adapted random measures such that $(\mu(\cdot, [0, t]))_{t \in [0, T]}$ and $(\nu(\cdot, [0, t]))_{t \in [0, T]}$ are continuous processes and $\mu(\omega, \{0\}) > 0$ for all $\omega \in A$. Let $p \in (1, \infty)$ and assume that

$$\mathbb{E} \left| \int_{[0,T]} \mu([0,t])^{\frac{1}{2}} d\nu(t) \right|^p < \infty.$$

Then we have that

$$(21) \quad \left\| \int_{[0,T]} \mu([0,t])^{-\frac{1}{2}} d\mu(t) \right\|_p \left\| \int_{[0,T]} \mu([0,t])^{\frac{1}{2}} d\nu(t) \right\|_p \leq 2p \|\mu([0,T])\|_{\frac{p}{2}} \|\nu\|_{\text{BMO}}.$$

PROOF. For $\theta = 1/2$ Lemma 5.14 gives that

$$\left\| \int_{[0,T]} \mu([0,t])^{-\frac{1}{2}} d\mu(t) \right\|_p \leq 2 \left[\mathbb{E} \mu([0,T])^{\frac{p}{2}} \right]^{\frac{1}{p}} = 2 \sqrt{\|\mu([0,T])\|_{\frac{p}{2}}}.$$

Moreover, by Lemma 5.19 applied to $\eta = 1/2$,

$$\left\| \int_{[0,T]} \mu([0,t])^{\frac{1}{2}} d\nu(t) \right\|_p \leq p \|\mu([0,T])^{\frac{1}{2}}\|_p \|\nu\|_{\text{BMO}} = p \sqrt{\|\mu([0,T])\|_{\frac{p}{2}}} \|\nu\|_{\text{BMO}}. \quad \square$$

COROLLARY 5.21. Let $(A_t)_{t \in [0, T]}$ and $(B_t)_{t \in [0, T]}$ be progressively measurable \mathbb{R} -valued processes such that $\mathbb{E} \int_0^T |B_t|^2 dt < \infty$ and $p \in [1, \infty)$. Then one has that

$$(22) \quad \left\| \int_0^T |A_t B_t| dt \right\|_p \leq c_{(5.21), p} \|A\|_{\mathbb{H}_p(S_2)} \|B\|_{\text{BMO}(S_2)}$$

with $c_{(5.21), p} := \sqrt{2p}$.

PROOF. We first assume that there is a $c > 0$ such that $|A_s(\omega)| \leq c$ and $|B_s(\omega)| \leq c$ for all $(s, \omega) \in [0, T] \times A$. For $\varepsilon > 0$ and the Dirac measure δ_0 in 0 define

$$d\mu_\varepsilon(t) := \varepsilon d\delta_0(t) + A_t^2 dt \quad \text{and} \quad d\nu(t) := B_t^2 dt.$$

Then, by Theorem 5.20,

$$\begin{aligned} & \left\| \int_0^T |A_t B_t| dt \right\|_p \\ & \leq \left\| \sqrt{\int_0^T \left| \varepsilon + \int_0^t A_s^2 ds \right|^{-\frac{1}{2}} |A_t|^2 dt} \right\|_{2p} \left\| \sqrt{\int_0^T \left| \varepsilon + \int_0^t A_s^2 ds \right|^{\frac{1}{2}} |B_t|^2 dt} \right\|_{2p} \\ & \leq \left\| \int_{[0, T]} |\mu_\varepsilon([0, t])|^{-\frac{1}{2}} d\mu_\varepsilon(t) \right\|_p^{\frac{1}{2}} \left\| \int_{[0, T]} |\mu_\varepsilon([0, t])|^{\frac{1}{2}} d\nu(t) \right\|_p^{\frac{1}{2}} \\ & \leq \sqrt{2p \|\mu_\varepsilon([0, T])\|_{\frac{p}{2}} \|\nu\|_{\text{BMO}}} \\ & = \sqrt{2p} \left\| \left(\varepsilon + \int_0^T |A_t|^2 dt \right)^{\frac{1}{2}} \right\|_p \sup_{t \in [0, T]} \left\| \mathbb{E} \left(\int_t^T |B_s|^2 ds | \mathcal{A}_t \right) \right\|_\infty^{\frac{1}{2}}. \end{aligned}$$

By $\varepsilon \downarrow 0$ we get that

$$\left\| \int_0^T |A_s B_s| ds \right\|_p \leq \sqrt{2p} \left\| \left(\int_0^T |A_t|^2 dt \right)^{\frac{1}{2}} \right\|_p \sup_{t \in [0, T]} \left\| \mathbb{E} \left(\int_t^T |B_s|^2 ds | \mathcal{A}_t \right) \right\|_\infty^{\frac{1}{2}}$$

whenever $|A_s(\omega)| \leq c$ and $|B_s(\omega)| \leq c$ for all $(s, \omega) \in [0, T] \times A$. By monotone convergence we can omit the restriction on A first, and finally we can do so for B as well. \square

REMARK 5.22. There is a connection to the *Bhattacharyya coefficient* (also called *Hellinger coefficient*) of two measures, see [10]. Assume two Borel measures μ, ν on $\mathcal{B}([0, T])$ and a reference measure σ such that μ and ν are absolutely continuous with respect to σ . Then

$$B(\mu, \nu) := \int_{[0, T]} \sqrt{\frac{d\mu}{d\sigma} \frac{d\nu}{d\sigma}} d\sigma,$$

which is independent from the particular choice of the reference measure, is called Bhattacharyya coefficient. Under the assumptions of Corollary 5.21, with $d\mu(t) := A_t^2 dt$ and $d\nu(t) := B_t^2 dt$, we have

$$B(\mu(\omega, \cdot), \nu(\omega, \cdot)) = \int_0^T |A_s(\omega) B_s(\omega)| ds.$$

COROLLARY 5.23. For $\theta \in (0, 1]$, $p \in [1, \infty)$, and $Z \in \mathbb{H}_p(S_2) \cap \text{BMO}(S_{2\theta})$ one has

$$\mathbb{E} \left| \int_0^T |Z_t|^{1+\theta} dt \right|^p < \infty$$

with

$$\left\| \int_0^T |Z_t|^{1+\theta} dt \right\|_p \leq c_{(5.21),p} \|Z\|_{\mathbb{H}_p(S_2)} \|Z\|_{\text{BMO}(S_{2\theta})}^\theta.$$

- REMARK 5.24. (1) For $\theta = 1$ we have that $\text{BMO}(S_{2\theta}) \subseteq \mathbb{H}_p(S_2)$ because of relation (20).
- (2) In general, for $\theta \in (0, 1)$ we do not have $\mathbb{H}_p(S_2) \subseteq \text{BMO}(S_{2\theta})$ (see Example 5.12) nor $\text{BMO}(S_{2\theta}) \subseteq \mathbb{H}_p(S_2)$ (here one can take deterministic processes).
- (3) In general, neither the condition $Z \in \mathbb{H}_p(S_2)$ implies $\mathbb{E}|\int_0^T |Z_s|^{1+\theta} ds|^p < \infty$ for $\theta \in (0, 1]$, nor $Z \in \text{BMO}(S_{2\theta})$ does for $\theta \in (0, 1)$.

CHAPTER 6

Applications to BSDEs

In this chapter we consider a solution to the BSDE

$$(23) \quad Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad t \in [0, T], \text{ a.s.},$$

and will proceed as follows: Firstly, we extend equation (23) from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\overline{\Omega}, \mathcal{F}^0, \overline{\mathbb{P}})$ and follow Chapter 3 to transform this extended BSDE from $(\overline{\Omega}, \mathcal{F}^0, \overline{\mathbb{P}})$ to $(\overline{\Omega}, \mathcal{F}^\varphi, \overline{\mathbb{P}})$ and $(\overline{\Omega}, \mathcal{F}^\psi, \overline{\mathbb{P}})$, respectively, and consider for $\rho \in \{\varphi, \psi\}$ the two solutions

$$(24) \quad Y_t^\rho = \xi^\rho + \int_t^T f^\rho(s, Y_s^\rho, Z_s^\rho) ds - \int_t^T Z_s^\rho dW_s^\rho, \quad t \in [0, T], \text{ a.s.}$$

Therefore (24) describes two copies of (23), parametrised with φ and ψ , by transforming the underlying Gaussian structure. Secondly, we interpret (24) as equations driven by the *joint Brownian motion* $\overline{W} = (\overline{W}_t)_{t \in [0, T]}$ and apply an a priori estimate to obtain Theorem 6.4 to describe the stability of (23). From the stability we obtain non-linear embeddings for Besov spaces in Section 6.4.4 and upper bounds for the L_p -variation of solution processes (Y, Z) to our BSDE (23) in Section 6.5.

6.1. The setting

In this section we assume a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0, T]})$ with $\mathcal{F} = \mathcal{F}_T$ satisfying the usual conditions, where $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ is the augmentation of the natural filtration of the d -dimensional Brownian motion $(W_t)_{t \in [0, T]}$. We consider a solution to the BSDE (23) under the following set of assumptions, that describe the generators we will use and ensure that all expressions do exist:

ASSUMPTION 6.1.

(B1) The process Z is predictable such that

$$\mathbb{P} \left(\int_0^T |Z_s|^2 ds < \infty \right) = 1.$$

(B2) The process Y is adapted and path-wise continuous.

(B3) The generator $f : \Omega_T \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ is such that $(t, \omega) \mapsto f(t, \omega, y, z)$ is predictable for all (y, z) and there are $L_Y, L_Z \geq 0$ and $\theta \in [0, 1]$ such that

$$|f(t, \omega, y_0, z_0) - f(t, \omega, y_1, z_1)| \leq L_Y |y_0 - y_1| + L_Z [1 + |z_0| + |z_1|]^\theta |z_0 - z_1|$$

for all $(t, \omega, y_0, y_1, z_0, z_1)$.

(B4) $\mathbb{P} \left(\int_0^T |f(s, Y_s, Z_s)| ds < \infty \right) = 1.$

The case $\theta = 0$ is the standard Lipschitz case, the case $\theta = 1$ the standard quadratic case, and $\theta \in (0, 1)$ can be seen as sub-quadratic case (see for example [25]). Our strategy for the first step is to impose in Lemma 6.2 below conditions on the gradient process Z and $f(s, 0, 0)$, only, but not on ξ , in order to verify that we deal with an L_p -solution to our BSDE. This might also help to find more general conditions on (ξ, f) that ensure the existence of L_p -solutions (see Section 6.4.1 below). Our conditions on Z can be verified by results from Section 6.3 below. In the following we assume that $p \in [2, \infty)$ because this assumption will be used in some steps of the proofs and because this case is more interesting with respect to the tail-behavior of $|Y_t - Y_s|$ than the case $p < 2$.

LEMMA 6.2. *In addition to the conditions (B1)-(B4) we assume for $p \in [2, \infty)$ that*

$$(B5) \quad \int_0^T |f(s, 0, 0)| ds \in \mathcal{L}_p,$$

$$(B6) \quad \left(\int_0^T |Z_s|^2 ds \right)^{\frac{1}{2}} \in \mathcal{L}_p,$$

$$(B7) \quad \int_0^T |Z_s|^{1+\theta} ds \in \mathcal{L}_p.$$

Then $\int_0^T |f(s, Y_s, Z_s)| ds + \sup_{t \in [0, T]} |Y_t| \in \mathcal{L}_p$.

PROOF. We rewrite (23) as

$$Y_t = Y_0 - \int_0^t f(s, Y_s, Z_s) ds + \int_0^t Z_s dW_s$$

for $t \in [0, T]$. For an integer $N \geq 1$ let

$$\tau_N := \inf \{t \in [0, T] : |Y_t - Y_0| = N\} \wedge T$$

with $\inf \emptyset := \infty$. Then

$$Y_{t \wedge \tau_N} = Y_0 - \int_0^{t \wedge \tau_N} f(s, Y_s, Z_s) ds + \int_0^{t \wedge \tau_N} Z_s dW_s.$$

Because of

$$(25) \quad |f(s, y, z)| \leq |f(s, 0, 0)| + L_y |y| + L_z [1 + |z|]^\theta |z|$$

we conclude that

$$\begin{aligned} |Y_{t \wedge \tau_N}| &\leq \left[|Y_0| + \int_0^T |f(s, 0, 0)| ds + L_z \int_0^T [1 + |Z_s|]^\theta |Z_s| ds \right. \\ &\quad \left. + \sup_{r \in [0, T]} \left| \int_0^r Z_s dW_s \right| \right] + L_y \int_0^{t \wedge \tau_N} |Y_{s \wedge \tau_N}| ds \\ &=: A + L_y \int_0^{t \wedge \tau_N} |Y_{s \wedge \tau_N}| ds \end{aligned}$$

and

$$M_t^N \leq A + L_y \int_0^t M_s^N ds$$

with

$$M_s^N := \sup_{r \in [0, s]} |Y_{r \wedge \tau_N}| = \sup_{r \in [0, s \wedge \tau_N]} |Y_r|.$$

The process $(M_t^N)_{t \in [0, T]}$ is continuous, adapted and bounded by $|Y_0| + N$. The inequality

$$\|M_t^N\|_p \leq \|A\|_p + L_y \int_0^t \|M_s^N\|_p ds$$

implies by Gronwall's lemma that

$$\|M_T^N\|_p \leq e^{L_y T} \|A\|_p.$$

Letting $N \rightarrow \infty$ gives $\sup_{t \in [0, T]} |Y_t| \in \mathcal{L}_p$ because $A \in \mathcal{L}_p$ which follows from conditions (B5), (B6), and (B7). Finally, using (25) the part $\int_0^T |f(s, Y_s, Z_s)| ds \in \mathcal{L}_p$ follows. \square

Condition (B5) is a condition on the initial data of the BSDE, whereas (B6) and (B7) are implicit conditions on the solution. For $\theta = 0$ condition (B6) implies (B7). Conversely, for $\theta = 1$ condition (B7) implies (B6). A sufficient condition for both, (B6) and (B7), is $\left(\int_0^T |Z_s|^2 ds\right)^{1/2} \in \mathcal{L}_{(1+\theta)p}$.

6.2. Stability of BSDEs with respect to perturbations of the Gaussian structure

Now we substantiate the procedure explained in the beginning of this chapter: we assume the setting of Section 4.2 and follow Convention 4.5(1) to extend (23) to $\overline{\Omega}$ and find

$$(26) \quad \tilde{Y}_t = \tilde{\xi} + \int_t^T \tilde{f}(s, \tilde{Y}_s, \tilde{Z}_s) ds - \int_t^T \tilde{Z}_s dW_s^0, \quad t \in [0, T].$$

We remark that for a $(\mathcal{P}, \mathcal{B}(C(M)))$ -measurable $h : [0, T] \times \Omega \rightarrow C(M)$ the extension $\tilde{h} : [0, T] \times \overline{\Omega} \rightarrow C(M)$ is $(\mathcal{P}^0, \mathcal{B}(C(M)))$ -measurable, and that there is a $\overline{\Omega}_0 \in \overline{\mathcal{F}}$ with $\overline{\mathbb{P}}(\overline{\Omega}_0) = 1$, such that $(\int_0^t \tilde{Z}_s dW_s^0)(\omega, \omega') = (\int_0^t Z_s dW_s)(\omega)$ for $t \in [0, T]$ and $(\omega, \omega') \in \overline{\Omega}_0$. Moreover, it is clear that the inequality from (B3) transfers directly. Therefore we assume that (23) is extended to (26) where we simplify the notation by denoting $(\tilde{\xi}, \tilde{f}, \tilde{Y}, \tilde{Z})$ again by (ξ, f, Y, Z) . Using Theorem 3.3 in the setting of Section 4.2 we obtain (24). We also know that the transformed generator f^ρ can be taken such that (B3) is satisfied, i.e.

$$\begin{aligned} & |f^\rho(t, \bar{\omega}, y_0, z_0) - f^\rho(t, \bar{\omega}, y_1, z_1)| \\ & \leq L_Y |y_0 - y_1| + L_Z [1 + |z_0| + |z_1|]^\theta |z_0 - z_1| =: H((y_0, z_0), (y_1, z_1)), \end{aligned}$$

which follows from Remark 3.4. Finally, we need a modified version of the sliceable numbers:

DEFINITION 6.3. Given a stochastic basis $(A, \mathcal{A}, \mathbb{Q}, \mathbb{A})$ with $\mathbb{A} = (\mathcal{A}_t)_{t \in [0, T]}$ and $\mathcal{A} = \mathcal{A}_T$ as in Chapter 5, an \mathbb{R} -valued progressively measurable process $c = (c_t)_{t \in [0, T]}$, and $N \geq 1$, we let

$$\text{sl}_N^{S_2}(c) = \text{sl}_N^{S_2, \mathbb{A}}(c) := \inf \varepsilon,$$

where the infimum is taken over all $\varepsilon > 0$ such that there are stopping times $0 = \tau_0 \leq \tau_1 \leq \dots \leq \tau_N = T$ with

$$\sup_{k=1, \dots, N} \|(\chi_{(\tau_{k-1}, \tau_k]}(t) c_t)_{t \in [0, T]}\|_{\text{BMO}(S_2)} \leq \varepsilon.$$

Now let us turn to our basic result. Our strategy is to impose the conditions (B1)-(B6) and an extra condition on Z on equation (23) in the context of the stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0, T]})$ we did start from, and then to deduce by Lemma 6.34 below the moment estimates in the extended setting of $(\overline{\Omega}, \overline{\mathbb{P}})$.

THEOREM 6.4. Assume $\theta \in [0, 1]$, for equation (23) conditions (B1)-(B4), and additionally $|Z| \in \text{BMO}(S_{2\theta})$ in the case $\theta \in (0, 1]$. Suppose that there is a non-increasing sequence $s = (s_N)_{N \geq 1} \subseteq [0, \infty)$, where $s_\infty := \lim_N s_N$, such that

$$\text{sl}_N^{S_2, \mathbb{F}}(|Z|^\theta) \leq s_N.$$

Suppose that conditions (B5)-(B6) are satisfied for $p \in [2, \infty)$ where in the case $s_\infty > 0$ we additionally assume that

$$p > p_0 := \frac{\Phi^{-1}(2\sqrt{2}L_Z s_\infty)}{\Phi^{-1}(2\sqrt{2}L_Z s_\infty) - 1} \in (1, \infty)$$

with the function Φ defined in (18). Then, one has for the extended equations for all $t \in [0, T]$ that

$$(27) \quad \left\| \sup_{s \in [t, T]} |Y_s^\varphi - Y_s^\psi| \right\|_p + \left\| \left(\int_t^T D[\varphi(s), \psi(s)] |Z_s|^2 ds \right)^{\frac{1}{2}} \right\|_p + \left\| \left(\int_t^T |Z_s^\varphi - Z_s^\psi|^2 ds \right)^{\frac{1}{2}} \right\|_p \leq c_{(6.4)} \left[\|\xi^\varphi - \xi^\psi\|_p + \left\| \int_t^T |f^\varphi(s, Y_s^\psi, Z_s^\psi) - f^\psi(s, Y_s^\psi, Z_s^\psi)| ds \right\|_p \right],$$

where $\varphi, \psi \in \Delta$, $D[\eta_1, \eta_2] := 1 - \sqrt{1 - \eta_1^2} \sqrt{1 - \eta_2^2} - \eta_1 \eta_2$, and $c_{(6.4)} > 0$ depends at most on $(L_Y, L_Z, T, (s_N)_{N=1}^\infty, p, d)$.

The applications of Theorem 6.4 are at least two-fold: Firstly, we obtain a non-linear embedding theorem for Besov spaces in Section 6.4.4 (Corollary 6.21). Secondly, we deduce in Section 6.5 upper bounds for the L_p -variation of solution processes (Y, Z) to our BSDE (23).

REMARK 6.5.

- (1) The function $D[\eta_1, \eta_2] : [0, 1]^2 \rightarrow [0, 1]$ measures the distance between η_1 and η_2 , by projecting the vector $(\eta_1, \sqrt{1 - \eta_1^2})$ onto the linear subspace generated by $(\eta_2, \sqrt{1 - \eta_2^2})$, and by comparing the projection to $(\eta_2, \sqrt{1 - \eta_2^2})$. In particular, $D[\eta_1, \eta_2] = 0$ if and only if $\eta_1 = \eta_2$.
- (2) Because the case $\lim_N s_N = 0$ is of particular importance in Theorem 6.4, as it enables us to use the full range $p \in [2, \infty)$, we give some examples for this situation:
 - (a) For $\theta = 0$ we have that

$$\text{sl}_N^{S_2, \mathbb{F}}(|Z|^\theta) \leq \sqrt{\frac{T}{N}}$$

if we take equidistant time-nets.

- (b) Let $0 < \theta < \eta \leq 1$ and assume that $\|Z\|_{\text{BMO}(S_{2\eta})} < \infty$. Then, similarly to Example 5.6, we obtain

$$\|(\chi_{(a,b]}(t)|Z_t|)_{t \in [0, T]}\|_{\text{BMO}(S_{2\theta})} \leq (b - a)^{\frac{1}{2\theta} - \frac{1}{2\eta}} \|(\chi_{(a,b]}(t)|Z_t|)_{t \in [0, T]}\|_{\text{BMO}(S_{2\eta})}$$

and, by using equidistant grids, that

$$\mathrm{sl}_N^{S_2, \mathbb{F}}(|Z|^\theta) \leq \left(\frac{T}{N}\right)^{\frac{1}{2}(1-\frac{\theta}{\eta})} \|Z\|_{\mathrm{BMO}(S_{2\eta})}^\theta.$$

- (3) The usage of $(\mathrm{sl}_N^{S_2, \mathbb{F}}(|Z|^\theta))_{N \geq 1}$ might not be optimal in extremal cases as we mainly need the reverse Hölder inequality for the Doléan-Dade exponential (34) in the proof of Theorem 6.4 below: If one would have $\int_0^\cdot c_s d\overline{W}_s \in \overline{L_\infty}^{\mathrm{BMO}_2}$, then according to the remarks following Proposition 5.8 the reverse Hölder inequality for all exponents would be satisfied. It is part of future work to check conditions on the gradient Z which guarantee this. On the other hand, if $\int_0^\cdot c_s d\overline{W}_s \notin \overline{L_\infty}^{\mathrm{BMO}_2}$, then our approach yields explicit bounds for $c_{(6.4)} > 0$ and the threshold p_0 in terms of $(s_N)_{N \geq 1}$ which is implicitly a novelty of this statement. As shown in Section 6.3 below, the usage of the sliceable numbers gives $s_\infty = 0$ in our relevant cases.

PROOF OF THEOREM 6.4. (a) By Corollary 5.23 the assumptions (B6) and $\|Z\|_{\mathrm{BMO}(S_{2\theta})} < \infty$ imply (B7) in the case $\theta > 0$, whereas for $\theta = 0$ condition (B6) implies (B7) directly. Therefore we have

$$(28) \quad \int_0^T |f(s, Y_s, Z_s)| ds + \sup_{t \in [0, T]} |Y_t| \in \mathcal{L}_p$$

by Lemma 6.2 for equation (23). This yields the validity of conditions (B1)-(B7) and (28) for the canonical extension to $\overline{\Omega}$.

(b) Now we define $h_1, h_2 : [0, 1]^2 \rightarrow [0, 1]$ by

$$\begin{aligned} h_1(x, z) &:= \frac{x\sqrt{1-z^2} + z\sqrt{1-x^2}}{x+z}, \\ h_2(x, z) &:= \frac{x\sqrt{1-z^2} + z\sqrt{1-x^2}}{\sqrt{1-z^2} + \sqrt{1-x^2}}, \end{aligned}$$

where for $x = z = 0$ we set $h_1 := 1$ and $h_2 := 0$, analogously for $x = z = 1$ we set $h_1 := 0$ and $h_2 := 1$, so that

$$\begin{pmatrix} \sqrt{1-x^2} & x \\ \sqrt{1-z^2} & z \end{pmatrix} \begin{pmatrix} h_1(x, z) \\ h_2(x, z) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

for all $x, z \in [0, 1]$. For $\rho \in \{\varphi, \psi\}$ we let

$$\begin{aligned} \overline{Z}_s^\rho &:= (Z_s^\rho \sqrt{1-\rho^2(s)}, Z_s^\rho \rho(s)), \\ \overline{f}^\rho(s, y, (z, z')) &:= f^\rho(s, y, h_1(\varphi(s), \psi(s))z + h_2(\varphi(s), \psi(s))z'), \end{aligned}$$

which leads to $\overline{f}^\rho(s, Y_s^\rho, \overline{Z}_s^\rho) = f^\rho(s, Y_s^\rho, Z_s^\rho)$ and

$$(29) \quad Y_t^\rho = \xi^\rho + \int_t^T \overline{f}^\rho(s, Y_s^\rho, \overline{Z}_s^\rho) ds - \int_t^T \overline{Z}_s^\rho d\overline{W}_s.$$

Observe that

$$\begin{aligned} &|\overline{Z}_s^\varphi - \overline{Z}_s^\psi|^2 \\ (30) \quad &= D[\varphi(s), \psi(s)][|Z_s^\psi|^2 + |Z_s^\varphi|^2] + [1 - D[\varphi(s), \psi(s)]]|Z_s^\psi - Z_s^\varphi|^2 \end{aligned}$$

$$\begin{aligned}
&\geq \frac{2 - D[\varphi(s), \psi(s)]}{2} |Z_s^\psi - Z_s^\varphi|^2 \\
(31) \quad &\geq \frac{|Z_s^\psi - Z_s^\varphi|^2}{2}
\end{aligned}$$

and therefore we get for

$$\begin{aligned}
c_s &:= \frac{\bar{f}^\varphi(s, Y_s^\varphi, \bar{Z}_s^\varphi) - \bar{f}^\varphi(s, Y_s^\varphi, \bar{Z}_s^\psi)}{|\bar{Z}_s^\varphi - \bar{Z}_s^\psi|^2} \chi_{\{\bar{Z}_s^\varphi \neq \bar{Z}_s^\psi\}} [\bar{Z}_s^\varphi - \bar{Z}_s^\psi] \\
&= \frac{f^\varphi(s, Y_s^\varphi, Z_s^\varphi) - f^\varphi(s, Y_s^\varphi, Z_s^\psi)}{|\bar{Z}_s^\varphi - \bar{Z}_s^\psi|^2} \chi_{\{\bar{Z}_s^\varphi \neq \bar{Z}_s^\psi\}} [\bar{Z}_s^\varphi - \bar{Z}_s^\psi]
\end{aligned}$$

that

$$\begin{aligned}
|c_s| &\leq \sqrt{2} \frac{|f^\varphi(s, Y_s^\varphi, Z_s^\varphi) - f^\varphi(s, Y_s^\varphi, Z_s^\psi)|}{|Z_s^\varphi - Z_s^\psi|} \chi_{\{Z_s^\varphi \neq Z_s^\psi\}} \\
&\leq \sqrt{2} L_Z [1 + |Z_s^\psi| + |Z_s^\varphi|]^\theta \\
&\leq \sqrt{2} L_Z [1 + |Z_s^\psi|^\theta + |Z_s^\varphi|^\theta].
\end{aligned}$$

Lemma 5.5 (to come into the setting of Lemma 5.5 one can pass from an \mathbb{R} -valued progressively measurable process $\alpha = (\alpha_t)_{t \in [0, T]}$ with $\mathbb{E} \int_0^T |\alpha_t|^2 dt < \infty$ to a martingale by, for example, $M_t := \int_0^t \alpha_s d\bar{W}_{s,1}$) gives that

$$(32) \quad \text{sl}_{3N-2}^{S_2, \bar{\mathbb{F}}}(|c|) \leq \sqrt{2} L_Z [\text{sl}_N^{S_2, \bar{\mathbb{F}}}(1) + \text{sl}_N^{S_2, \bar{\mathbb{F}}}(|Z^\psi|^\theta) + \text{sl}_N^{S_2, \bar{\mathbb{F}}}(|Z^\varphi|^\theta)].$$

(c) We return to the stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0, T]})$, take $\eta > 0$ and find a sequence of stopping times $0 = \tau_0 \leq \dots \leq \tau_N = T$ such that

$$\sup_{k=1, \dots, N} \|(\chi_{(\tau_{k-1}, \tau_k]}(t) |Z_t|^\theta)_{t \in [0, T]}\|_{\text{BMO}(S_2)} \leq \text{sl}_N^{S_2, \bar{\mathbb{F}}}(|Z|^\theta) + \eta \leq s_N + \eta.$$

Letting

$$Z_t^k := \chi_{(\tau_{k-1}, \tau_k]}(t) Z_t,$$

one can quickly check that

$$\mathbb{E} \left(\int_t^T |Z_s^k|^{2\theta} ds | \mathcal{F}_t^0 \right) \leq (s_N + \eta)^2$$

for all deterministic $t \in [0, T]$, where Z^k is canonically extended to $\bar{\Omega}$. Assuming an $(\mathcal{F}_t^0)_{t \in [0, T]}$ -stopping time $\tau : \bar{\Omega} \rightarrow [0, T]$, and using the decomposition

$$\mathbb{E} \left(\int_\tau^T |Z_s^k|^{2\theta} ds | \mathcal{F}_\tau^0 \right) = \mathbb{E} \left(\int_0^T |Z_s^k|^{2\theta} ds | \mathcal{F}_\tau^0 \right) - \int_0^\tau |Z_s^k|^{2\theta} ds$$

and the optional stopping theorem, we may deduce that

$$\mathbb{E} \left(\int_\tau^T |Z_s^k|^{2\theta} ds | \mathcal{F}_\tau^0 \right) \leq (s_N + \eta)^2.$$

Consequently,

$$\sup_{k=1, \dots, N} \|(\chi_{(\tau_{k-1}, \tau_k]}(t) |Z_t|^\theta)_{t \in [0, T]}\|_{\text{BMO}(S_2)} \leq s_N + \eta$$

also after extending Z and $(\tau_k)_{k=0}^N$ to $\bar{\Omega}$ where the filtration $(\mathcal{F}_t^0)_{t \in [0, T]}$ is used. This means that

$$(33) \quad \text{sl}_N^{S_2, \mathbb{F}^0}(|Z|^\theta) \leq s_N.$$

(d) For any stopping time $\tau : \Omega \rightarrow [0, T]$ relative to $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0, T]})$ and for $\rho \in \{\psi, \varphi\}$ consider $\tau^\rho : \bar{\Omega} \rightarrow \mathbb{R}$ and take a representative such that $\tau^\rho : \bar{\Omega} \rightarrow [0, T]$. It is easy to check that τ^ρ is a stopping time with respect to the filtration $(\mathcal{F}_t^\rho)_{t \in [0, T]}$. Using $\mathbb{E}(A^\rho | \bar{\mathcal{F}}_t) = (\mathbb{E}(A | \mathcal{F}_t^0))^\rho \bar{\mathbb{P}}\text{-a.s.}$ for $A \in \mathcal{L}_1(\bar{\Omega}, \mathcal{F}^0, \bar{\mathbb{P}})$ (which can be checked by taking simple A that depend only on finitely many increments of the Brownian motion W and then passing in L_1 to the limit), Theorem 2.6, and Remark 2.7(2) yield that

$$\int_t^T \chi_{(\tau_{k-1}^\rho, \tau_k^\rho]}(s) |Z_s^\rho|^{2\theta} ds = \left(\int_t^T \chi_{(\tau_{k-1}, \tau_k]}(s) |Z_s|^{2\theta} ds \right)^\rho \bar{\mathbb{P}}\text{-a.s.}$$

and

$$\begin{aligned} \mathbb{E} \left(\int_t^T \chi_{(\tau_{k-1}^\rho, \tau_k^\rho]}(s) |Z_s^\rho|^{2\theta} ds | \bar{\mathcal{F}}_t \right) &= \left(\mathbb{E} \left(\int_t^T \chi_{(\tau_{k-1}, \tau_k]}(s) |Z_s|^{2\theta} ds | \mathcal{F}_t^0 \right) \right)^\rho \\ &\leq (s_N + \eta)^2. \end{aligned}$$

Therefore, we obtain $\text{sl}_N^{S_2, \bar{\mathbb{F}}}(|Z|^\theta) \leq s_N + \eta$ as a complement of (33) (where we use the same optional stopping argument as in step (c)) and can continue from (32) to

$$\text{sl}_{3N-2}^{S_2, \bar{\mathbb{F}}}(|c|) \leq \sqrt{2}L_Z \left[\sqrt{\frac{T}{N}} + 2s_N + 2\eta \right]$$

and

$$\text{sl}_{3N-2}^{S_2, \bar{\mathbb{F}}}(|c|) \leq \sqrt{2}L_Z \left[\sqrt{\frac{T}{N}} + 2s_N \right]$$

by $\eta \downarrow 0$. In the case $s_\infty = 0$ take $p_0 \in (1, 2)$, say $p_0 := 3/2$, and in the case $s_\infty > 0$, define

$$p_0 := \frac{\Phi^{-1}(2\sqrt{2}L_Z s_\infty)}{\Phi^{-1}(2\sqrt{2}L_Z s_\infty) - 1} \in (1, \infty)$$

and $p_1 := (p + p_0)/2$ so that

$$1 < p_0 < p_1 < p < \infty.$$

Let

$$(34) \quad \lambda_t := \exp \left(\int_0^t c_s d\bar{W}_s - \frac{1}{2} \int_0^t |c_s|^2 ds \right).$$

We find an $N \geq 1$ such that

$$\text{sl}_{3N-2}^{S_2, \bar{\mathbb{F}}}(|c|) \leq \sqrt{2}L_Z \left[\sqrt{\frac{T}{N}} + 2s_N \right] < \Phi(p'_1).$$

This N depends at most on $((s_N)_{N=1}^\infty, L_Z, T, p)$. Theorem 5.9 implies that

$$\mathcal{RH}_{p'_1}(\lambda) \leq \left[\Psi \left(\sqrt{2}L_Z \left[\sqrt{\frac{T}{N}} + 2s_N \right], p'_1 \right) \right]^{3N-2} < \infty$$

with Ψ taken from (19). By assumption (B6) we have that

$$\left(\int_0^T |Z_s|^2 ds \right)^{\frac{1}{2}} \in \mathcal{L}_p.$$

Finally, fixing $t \in [0, T]$, we can assume for this t that

$$\left\| \int_t^T |f^\varphi(s, Y_s^\psi, Z_s^\psi) - f^\psi(s, Y_s^\psi, Z_s^\psi)| ds \right\|_p < \infty,$$

otherwise there is nothing to prove. So we can apply Lemma 6.34 below to the equations (29) for $\rho \in \{\varphi, \psi\}$ and conclude by using (30) and (31). \square

6.3. On classes of quadratic and sub-quadratic BSDEs

In this section we present results about particular classes of quadratic and sub-quadratic BSDEs that might be of independent interest. At the same time we check whether we may apply Theorem 6.4 to these BSDEs and what we can say about the critical value s_∞ .

There are various articles that describe the existence and quantitative properties of solutions to BSDEs and provide comparison results. For the case $\theta = 0$ the reader is referred to [18] and the references therein, and for the quadratic case we refer to [51, 52, 53, 43, 20, 2, 21, 22, 57, 46, 27, 56, 5, 28]. We are mainly interested in the sub-quadratic and quadratic case, i.e. the case when $\theta \in (0, 1]$. In Table 1 below we describe how we will embed these cases in the framework of this article. Table 1 should be read in the way that we first choose (ξ, θ, f) , then we obtain the integrability of the gradient process $Z = (Z_t)_{t \in [0, T]}$ and the conclusion for $s_\infty = \lim_N s_N$ from Theorem 6.4. In the cases where the uniqueness of the solution is not known there exists a solution with the stated properties. Finally, in (IV-V) we leave out the range for s_∞ as we do not have general results for these cases.

	ξ	θ	f	$ Z $	s_∞
I	$\xi \in L_p$ for some $p \in [2, \infty)$	0	(B3), (B5)	$\mathbb{H}_p(S_2)$	0
II	$\xi \in \text{cExp}$	$(0, 1)$	(B3), (B8)	$\mathbb{H}_2(S_2) \cap \bigcap_{\eta \in (0, 1)} \text{BMO}(S_{2\eta})$	0
III	$ \xi _{\text{cExp}(\eta, \mu)} < \infty$ for some $\eta \in (0, 1], \mu > \gamma e^{\beta T}$	1	(B3), (B8)	$\mathbb{H}_2(S_2) \cap \text{BMO}(S_{2\eta})$	$[0, \infty)$ if $\eta = 1$
IV	$\mathbb{E} e^{\mu \xi } < \infty$ for some $\mu > 0$	$(0, 1)$	(B3), (B8)	$\text{BMO}^{\sqrt{\Psi}}(S_2)$	
V	$\mathbb{E} e^{\mu \xi } < \infty$ for some $\mu > \gamma e^{\beta T}$	1	(B3), (B8)	$\text{BMO}^{\sqrt{\Psi}}(S_2)$	

Table 1:

We note that $|Z| \in \text{BMO}^{\sqrt{\Psi}}(S_2)$ also implies $|Z| \in \mathbb{H}_2(S_2)$. The case (I) follows from [18, Theorem 4.2] that gives (B6) and Remark 6.5(2a) yields to $s_\infty = 0$. In the following we verify our contribution (II)-(V).

Notation and setting. There is a series of papers dealing with the quadratic case where the terminal condition is unbounded, see [20, 21, 27, 28]. Below we use the setting of the initial article [20]. For future work some extensions of [20] done in [57] might be of interest for our context as well. To use the setting of [20] we introduce constants $\alpha \geq 0$ and $\beta, \gamma > 0$ such that, for all $\omega \in \Omega$,

$$(35) \quad |f(s, \omega, y, z)| \leq \alpha + \beta|y| + \frac{\gamma}{2}|z|^2 \quad \text{and} \quad \alpha \geq \frac{\beta}{\gamma}.$$

In our framework we can choose and fix these constants as

$$(36) \quad \alpha := \max \left\{ \sup_{(t, \omega) \in \Omega \times [0, T]} |f(t, \omega, 0, 0)| + L_Z, \frac{L_Y}{4L_Z} \right\},$$

$$(37) \quad \beta := L_Y,$$

$$(38) \quad \gamma := 4L_Z,$$

where we suppose, for the remainder of this section, condition

$$(B8) \quad \sup_{(t, \omega) \in \Omega \times [0, T]} |f(t, \omega, 0, 0)| < \infty, \quad L_Y > 0, \quad \text{and} \quad L_Z > 0.$$

As in [20] we use the function $\Phi_t : [0, \infty) \rightarrow (0, \infty)$ given by

$$\Phi_t(y) := e^{\gamma \alpha \frac{e^{\beta(T-t)} - 1}{\beta}} e^{\gamma y e^{\beta(T-t)}}.$$

Moreover, we set

$$\mu_T := \gamma e^{\beta T} > \gamma$$

which plays the role of a critical exponent in the case $\theta = 1$. Applying [20, Theorem 2] and inspecting its proof gives the following statement:

THEOREM 6.6 ([20]). *If there exists a $\mu > \mu_T$ such that*

$$\mathbb{E} e^{\mu|\xi|} < \infty,$$

then there is a solution to the BSDE (23) such that

- (1) $e^{\gamma|Y_t|} \leq \mathbb{E}(\Phi_t(|\xi|)|\mathcal{F}_t)$ a.s. for $t \in [0, T]$,
- (2) $|Z| \in \mathbb{H}_2(S_2)$,
- (3) for $0 \leq s < t \leq T$ and $\varepsilon > 0$ with $\gamma + \varepsilon < \mu$ one has

$$\mathbb{E} \left(\int_s^t |Z_r|^2 dr | \mathcal{F}_s \right) \leq c_{(6.6)}^2 \mathbb{E} \left(\sup_{r \in [s, t]} e^{(\gamma + \varepsilon)|Y_r|} | \mathcal{F}_s \right) \quad \text{a.s.}$$

$$\text{for } c_{(6.6)}^2 := 2 \left[\frac{1}{\gamma^2} + \frac{T}{\gamma} \max\left\{ \alpha, \frac{\beta}{\varepsilon} \right\} \right].$$

Verification of IV-V. Here our main observation consists in

THEOREM 6.7. *Let $\theta \in (0, 1]$ and assume that $\mu > \mu_T$ if $\theta = 1$ and $\mu > 0$ if $\theta \in (0, 1)$. If $\mathbb{E}e^{\mu|\xi|} < \infty$, then there is a solution to the BSDE (23) such that*

$$(39) \quad \mathbb{E} \left(\int_s^T |Z_r|^2 dr | \mathcal{F}_s \right) \leq c_{(6.7)}^2 \Psi_s \quad \text{with} \quad \Psi_s := \mathbb{E} \left(e^{\mu|\xi|} | \mathcal{F}_s \right)$$

for all $s \in [0, T]$ and $c_{(6.7)} = c(\alpha, \beta, \gamma, T, \theta, \mu) \in (0, \infty)$, where we may assume $(\Psi_s)_{s \in [0, T]}$ to be path-wise continuous. Moreover, for all stopping times $\tau : \Omega \rightarrow [0, T]$, $B \in \mathcal{F}_\tau$ of positive measure, and $\lambda, \nu > 0$, one has

$$\mathbb{P}_B \left(\int_\tau^T |Z_r|^2 dr > \lambda \nu \right) \leq e^{1-\lambda} + \delta \mathbb{P}_B \left(\sup_{s \in [\tau, T]} \Psi_s > \frac{\nu}{D} \right),$$

where \mathbb{P}_B is the normalized restriction of \mathbb{P} to B , $D = D(\alpha, \beta, \gamma, T, \theta, \mu) > 0$, and $\delta > 0$ is an absolute constant.

In the spirit of [37, Definition 1] the inequality (39) could be abbreviated by

$$\|Z\|_{\text{BMO} \sqrt{\Psi}(S_2)} \leq c_{(6.7)}.$$

PROOF OF THEOREM 6.7. Case $\theta = 1$: We choose $\varepsilon > 0$ and $p \in (1, \infty)$ such that

$$\mu = p\mu_T = \frac{\gamma + \varepsilon}{\gamma} \mu_T$$

which implies by $\beta > 0$ that $\gamma + \varepsilon < \mu$. Assuming $0 \leq s \leq T$ and applying Theorem 6.6 gives, a.s., that

$$\begin{aligned} \mathbb{E} \left(\int_s^T |Z_r|^2 dr | \mathcal{F}_s \right) &\leq c_{(6.6)}^2 \mathbb{E} \left(\sup_{r \in [s, T]} e^{(\gamma + \varepsilon)|Y_r|} | \mathcal{F}_s \right) \\ &= c_{(6.6)}^2 \mathbb{E} \left(\sup_{r \in [s, T]} e^{p\gamma|Y_r|} | \mathcal{F}_s \right) \\ &\leq c_{(6.6)}^2 \mathbb{E} \left(\sup_{r \in [s, T]} \left[\mathbb{E} \left(\Phi_s(|\xi|) | \mathcal{F}_r \right) \right]^p | \mathcal{F}_s \right) \\ &\leq c_{(6.6)}^2 \left| \frac{p}{p-1} \right|^p \mathbb{E} \left(\Phi_s(|\xi|)^p | \mathcal{F}_s \right) \\ &\leq c_{(6.6)}^2 \left| \frac{p}{p-1} \right|^p \kappa_T^p \mathbb{E} \left(e^{\mu|\xi|} | \mathcal{F}_s \right), \end{aligned}$$

where $\kappa_T := e^{\gamma\alpha\frac{\varepsilon\beta T}{\beta}-1}$ and for $(\mathbb{E}(\Phi_s(|\xi|)|\mathcal{F}_r))_{r \in [0, T]}$ a continuous modification is taken. Therefore, letting

$$c^2 = c^2(\alpha, \beta, \gamma, T, \mu) := c_{(6.6)}^2 \left| \kappa_T \frac{p}{p-1} \right|^p,$$

we proved

$$\mathbb{E} \left(\int_s^T |Z_r|^2 dr | \mathcal{F}_s \right) \leq c^2 \Psi_s \text{ a.s.}$$

Using an optional stopping argument, this can be extended to

$$\mathbb{E} \left(\int_{\tau}^T |Z_r|^2 dr | \mathcal{F}_{\tau} \right) \leq c^2 \Psi_{\tau} \text{ a.s.}$$

for any stopping time $\tau : \Omega \rightarrow [0, T]$. Given $\nu > 0$ we get

$$\begin{aligned} \mathbb{P}_B \left(\int_{\tau}^T |Z_r|^2 dr > 3\nu \right) &\leq \mathbb{P}_B \left(\int_{\tau}^T |Z_r|^2 dr > 3c^2 \Psi_{\tau} \right) + \mathbb{P}_B (c^2 \Psi_{\tau} > \nu) \\ &\leq \frac{1}{3} + \mathbb{P}_B (c^2 \Psi_{\tau} > \nu). \end{aligned}$$

If we define

$$W(B, \nu; \tau) := \mathbb{P} \left(B \cap \left\{ \sup_{r \in [\tau, T]} 3c^2 \Psi_r > \nu \right\} \right),$$

then we can directly apply [37, Theorem 1].

Case $\theta \in (0, 1)$: This case can be considered exactly as the case $\theta = 1$. In fact, with our choice of parameters (α, β, γ) in (36), (37), and (38) we obtain the estimate

$$|f(s, \omega, y, z)| \leq \alpha + \beta|y| + \frac{\gamma}{2}|z|^{1+\theta}.$$

But now, for any given $\tilde{\gamma} > 0$ we find an $\tilde{\alpha} \geq 0$ such that

$$\alpha + \frac{\gamma}{2}|z|^{1+\theta} \leq \tilde{\alpha} + \frac{\tilde{\gamma}}{2}|z|^2$$

for all $z \in \mathbb{R}^d$. In other words, we can arrange the parameters such that $\mu > \tilde{\gamma}e^{\beta T}$ (and have an additional dependence of the constants on θ). \square

Verification of II-III. The next definition will allow us to deduce that the gradient process Z belongs to $\text{BMO}(S_{2\eta})$:

DEFINITION 6.8. For $\eta \in (0, 1]$ and $\mu \in (0, \infty)$ we let

$$|\xi|_{\text{cExp}(\eta, \mu)} := \sup_{t \in [0, T]} (T - t)^{\frac{1}{\eta} - 1} \left\| \mathbb{E}(e^{\mu|\xi|} | \mathcal{F}_t) \right\|_{\infty}.$$

In the notation cExp above, 'c' stands for *conditional* and 'Exp' for *experiential*.

REMARK 6.9.

- (1) For $\eta = 1$ we have that $|\xi|_{\text{cExp}(1, \mu)} = e^{\mu\|\xi\|_{\infty}}$.
- (2) For $\xi \in L_2$, $0 < \eta < \tilde{\eta} < 1$, and $0 < \tilde{\mu} < \mu < \infty$ with $\mu(\frac{1}{\eta} - 1) = \tilde{\mu}(\frac{1}{\tilde{\eta}} - 1)$ one has $|\xi|_{\text{cExp}(\tilde{\eta}, \tilde{\mu})}^{\mu} \leq |\xi|_{\text{cExp}(\eta, \mu)}^{\tilde{\mu}}$.
- (3) For $\xi \in L_2$ and $\eta_0, \eta_1 \in (0, 1)$ one has $|\xi|_{\text{cExp}(\eta_0, \mu_0)} < \infty$ for some $\mu_0 \in (0, \infty)$ if and only if $|\xi|_{\text{cExp}(\eta_1, \mu_1)} < \infty$ for some $\mu_1 \in (0, \infty)$.

PROOF. Part (1) is obvious, (3) follows directly from (2). The assertion (2) is a consequence of

$$\begin{aligned} |\xi|_{\text{cExp}(\tilde{\eta}, \tilde{\mu})} &= \sup_{t \in [0, T]} (T - t)^{\frac{1}{\tilde{\eta}} - 1} \left\| \mathbb{E}(e^{\tilde{\mu}|\xi|} | \mathcal{F}_t) \right\|_{\infty} \\ &= \sup_{t \in [0, T]} (T - t)^{\frac{1}{\eta} - 1} \left\| \mathbb{E}(e^{\mu \frac{\tilde{\mu}}{\mu} |\xi|} | \mathcal{F}_t) \right\|_{\infty} \end{aligned}$$

$$\begin{aligned}
&\leq \sup_{t \in [0, T]} (T - t)^{\frac{1}{\eta} - 1} \left\| \mathbb{E}(e^{\mu|\xi|} | \mathcal{F}_t) \right\|_{\infty}^{\frac{\bar{\mu}}{\mu}} \\
&= \left[\sup_{t \in [0, T]} (T - t)^{\frac{1}{\eta} - 1} \left\| \mathbb{E}(e^{\mu|\xi|} | \mathcal{F}_t) \right\|_{\infty} \right]^{\frac{\bar{\mu}}{\mu}} \\
&= |\xi|_{\text{cExp}(\eta, \mu)}^{\frac{\bar{\mu}}{\mu}}.
\end{aligned}$$

□

Directly from Theorem 6.7 we deduce

COROLLARY 6.10. *Assume $\theta = 1$, $\eta \in (0, 1]$, and in addition to the assumptions made in Theorem 6.7 that $|\xi|_{\text{cExp}(\eta, \mu)} < \infty$ for some $\mu \in (0, \infty)$. Then $|Z| \in \text{BMO}(S_{2\eta})$ with*

$$\|Z\|_{\text{BMO}(S_{2\eta})} \leq c_{(6.7)} |\xi|_{\text{cExp}(\eta, \mu)}^{\frac{1}{2}}.$$

PROOF. We simply have that

$$\mathbb{E} \left(\int_s^T |Z_r|^2 dr | \mathcal{F}_s \right) \leq c_{(6.7)}^2 \mathbb{E} \left(e^{\mu|\xi|} | \mathcal{F}_s \right) \leq c_{(6.7)}^2 |\xi|_{\text{cExp}(\eta, \mu)} (T - s)^{1 - \frac{1}{\eta}} \text{ a.s.}$$

for all $s \in [0, T]$ and therefore, a.s.,

$$\mathbb{E} \left(\left(\int_s^T |Z_r|^2 dr \right)^{\frac{1}{\eta}} | \mathcal{F}_s \right) \leq (T - s)^{\frac{1}{\eta} - 1} \mathbb{E} \left(\int_s^T |Z_r|^2 dr | \mathcal{F}_s \right) \leq c_{(6.7)}^2 |\xi|_{\text{cExp}(\eta, \mu)}.$$

□

The above corollary explains the case (III) from Table 1. It turns out that in the remaining case (II) the particular choice of parameter η in $|\cdot|_{\text{cExp}(\eta, \mu)}$ does not have an impact. This is reflected by the following notation:

DEFINITION 6.11.

- (1) For a càdlàg process $Y = (Y_t)_{t \in [0, T]}$ and $(\eta, \mu) \in (0, 1) \times (0, \infty)$ we let

$$|Y|_{\text{cExp}(\eta, \mu)} := \sup_{t \in [0, T]} (T - t)^{\frac{1}{\eta} - 1} \left\| \mathbb{E}(e^{\mu \sup_{s \in [t, T]} |Y_s|} | \mathcal{F}_t) \right\|_{\infty}.$$

We say that $Y \in \text{cExp}$ provided that $|Y|_{\text{cExp}(\eta, \mu)} < \infty$ for some $(\eta, \mu) \in (0, 1) \times (0, \infty)$.

- (2) We say $\xi \in \text{cExp}$ provided that $|\xi|_{\text{cExp}(\eta, \mu)} < \infty$ for some $(\eta, \mu) \in (0, 1) \times (0, \infty)$.

The definition of $|Y|_{\text{cExp}(\eta, \mu)}$ is consistent with Definition 6.8 as for a random variable ξ we may let $Y_t := \xi$ and get $|Y|_{\text{cExp}(\eta, \mu)} = |\xi|_{\text{cExp}(\eta, \mu)}$.

REMARK 6.12. Exactly as in Remark 6.9 one can show that for $\eta_0, \eta_1 \in (0, 1)$ one has $|Y|_{\text{cExp}(\eta_0, \mu_0)} < \infty$ for some $\mu_0 \in (0, \infty)$ if and only if $|Y|_{\text{cExp}(\eta_1, \mu_1)} < \infty$ for some $\mu_1 \in (0, \infty)$. Therefore, $Y \in \text{cExp}$ if and only if there is some $\mu \in (0, \infty)$ such that

$$\sup_{t \in [0, T]} (T - t) \left\| \mathbb{E}(e^{\mu \sup_{s \in [t, T]} |Y_s|} | \mathcal{F}_t) \right\|_{\infty} < \infty.$$

THEOREM 6.13. *Assume that $\theta \in (0, 1)$ and $\xi \in \text{cExp}$. Then there is a unique solution (Y, Z) to the BSDE (23) in the class where $Y \in \text{cExp}$ and $|Z| \in \mathbb{H}_2(S_2)$. Moreover, for this solution we have that*

- (1) $s_\infty = 0$ for s_∞ defined as in Theorem 6.4,
- (2) $|Z| \in \text{BMO}(S_{2\eta})$ for all $\eta \in (0, 1)$.

For the uniqueness in the above theorem we do not assume convexity properties of the generator. Instead of that, we use $|Z| \in \text{BMO}(S_{2\theta})$ and follow the methodology that BMO-properties of the Z process give uniqueness, see for example [43]. The difference to previous settings is that we exploit that the generator is sub-quadratic and get therefore a weaker condition than the standard BMO-condition $|Z| \in \text{BMO}(S_2)$. Note that according to Example 5.12 the spaces $\text{BMO}(S_{2\eta})$ do not coincide for different $\eta \in (0, 1]$ in general.

PROOF OF THEOREM 6.13. Existence: The condition $\xi \in \text{cExp}$ implies that there are $(\eta, \mu) \in (0, 1) \times (0, \infty)$ such that

$$|\xi|_{\text{cExp}(\eta, \mu)} = \sup_{t \in [0, T)} (T - t)^{\frac{1}{\eta} - 1} \left\| \mathbb{E}(e^{\mu|\xi|} | \mathcal{F}_t) \right\|_\infty < \infty.$$

Because of $\theta < 1$ we use the argument for the case $\theta \in (0, 1)$ from the proof of Theorem 6.7 to replace (α, β, γ) by $(\tilde{\alpha}, \beta, \tilde{\gamma})$ such that

$$\mu > \tilde{\mu}_T := \tilde{\gamma} e^{\beta T} > \tilde{\gamma}.$$

We apply Theorem 6.6 and obtain a solution with

- (1) $e^{\tilde{\gamma}|Y_t|} \leq \mathbb{E}(\tilde{\Phi}_t(|\xi|) | \mathcal{F}_t)$ a.s. for $t \in [0, T]$,
- (2) $|Z| \in \mathbb{H}_2(S_2)$,

where $\tilde{\Phi}_t$ is defined as Φ_t with (α, β, γ) replaced by $(\tilde{\alpha}, \beta, \tilde{\gamma})$. Let $\tilde{p} := \mu / \tilde{\mu}_T \in (1, \infty)$ and assume $\tilde{\gamma} + \varepsilon < \mu$ for some $\varepsilon > 0$. Assuming $s \in [0, T)$, the arguments from the proof of Theorem 6.7 give, a.s., that

$$\begin{aligned} \mathbb{E} \left(\sup_{r \in [s, T]} e^{(\tilde{\gamma} + \varepsilon)|Y_r|} | \mathcal{F}_s \right) &\leq \left| \frac{\tilde{p}}{\tilde{p} - 1} \right|^{\tilde{p}} \tilde{\kappa}_T^{\tilde{p}} \mathbb{E} \left(e^{\mu|\xi|} | \mathcal{F}_s \right) \\ &\leq \left| \frac{\tilde{p}}{\tilde{p} - 1} \right|^{\tilde{p}} \tilde{\kappa}_T^{\tilde{p}} |\xi|_{\text{cExp}(\eta, \mu)} (T - s)^{1 - \frac{1}{\eta}} \end{aligned}$$

where $\tilde{\kappa}_T := e^{\tilde{\gamma}\tilde{\alpha} \frac{e^{\beta T} - 1}{\beta}}$. Therefore, $|Y|_{\text{cExp}(\eta, \tilde{\gamma} + \varepsilon)} < \infty$ and $Y \in \text{cExp}$.

Uniqueness: Assume two solutions (Y^0, Z^0) and (Y^1, Z^1) with $Y^0, Y^1 \in \text{cExp}$ and $Z^0, Z^1 \in \mathbb{H}_2(S_2)$. Let us fix $\eta \in (0, 1)$ and find $\mu_0, \mu_1 \in (0, \infty)$ such that

$$|Y^i|_{\text{cExp}(\eta, \mu_i)} = \sup_{t \in [0, T)} (T - t)^{\frac{1}{\eta} - 1} \left\| \mathbb{E}(e^{\mu_i \sup_{s \in [t, T]} |Y_s^i|} | \mathcal{F}_t) \right\|_\infty < \infty.$$

Again exploiting $\theta < 1$, we change in (35) the parameters (α, β, γ) to $(\tilde{\alpha}, \beta, \tilde{\gamma})$ such that

$$\mu := \min\{\mu_0, \mu_1\} > \tilde{\gamma} e^{\beta T}.$$

Analyzing the proof of [20, Theorem 2, pp. 609-610] gives for $0 \leq s < T$ and $\varepsilon > 0$ with $\tilde{\gamma} + \varepsilon < \mu$ that

$$\mathbb{E} \left(\int_s^T |Z_r^i|^2 dr | \mathcal{F}_s \right) \leq 2 \left[\frac{1}{\tilde{\gamma}^2} + \frac{T}{\tilde{\gamma}} \max \left\{ \tilde{\alpha}, \frac{\tilde{\beta}}{\varepsilon} \right\} \right] \mathbb{E} \left(\sup_{r \in [s, T]} e^{(\tilde{\gamma} + \varepsilon)|Y_r^i|} | \mathcal{F}_s \right) \text{ a.s.}$$

We continue with

$$\mathbb{E} \left(\sup_{r \in [s, t]} e^{(\tilde{\gamma} + \varepsilon)|Y_r^i|} | \mathcal{F}_s \right) \leq \mathbb{E} \left(\sup_{r \in [s, t]} e^{\mu_i |Y_r^i|} | \mathcal{F}_s \right) \leq |Y^i|_{\text{cExp}(\eta, \mu_i)} (T - s)^{1 - \frac{1}{\eta}}.$$

Therefore, for $\tilde{c}^2 := 2 \left[\frac{1}{\tilde{\gamma}^2} + \frac{T}{\tilde{\gamma}} \max \left\{ \tilde{\alpha}, \frac{\tilde{\beta}}{\varepsilon} \right\} \right]$, a.s.,

$$\mathbb{E} \left(\left(\int_s^T |Z_r^i|^{2\eta} dr \right)^{\frac{1}{\eta}} | \mathcal{F}_s \right) \leq (T - s)^{\frac{1}{\eta} - 1} \mathbb{E} \left(\int_s^T |Z_r^i|^2 dr | \mathcal{F}_s \right) \leq \tilde{c}^2 |Y^i|_{\text{cExp}(\eta, \mu_i)}.$$

This implies that $Z^0, Z^1 \in \text{BMO}(S_{2\eta})$ for all $\eta \in (0, 1)$. In particular, we have that $Z^0, Z^1 \in \text{BMO}(S_{2\theta})$ and this enables us to apply Lemma 6.34. Here we set

$$\begin{aligned} f^0(s, y, z) &:= f(s, y, z), \\ f^1(s) &:= f(s, Y_s^1, Z_s^1). \end{aligned}$$

The assumptions (D1), (D2), and (D4) are obviously satisfied, for (D3) we use that

$$\left(\mathbb{E} \left| \int_0^T |Z_s^i|^{1+\theta} ds \right|^2 \right)^{\frac{1}{2}} \leq c_{(5.21), 2} \|Z^i\|_{\mathbb{H}_2(S_2)} \|Z^i\|_{\text{BMO}(S_2)}^\theta$$

where $\|Z^i\|_{\text{BMO}(S_2)}^\theta < \infty$ because of $|Z^i| \in \text{BMO}(S_{2\theta})$. The above definitions guarantee that $\Xi_s \equiv 0$. A straightforward computation gives also that

$$\mathbb{E} \left(\int_t^T |c_s|^2 ds | \mathcal{F}_t \right) \leq L_Z^2 3^{2\theta} \left[T + \|Z^0\|_{\text{BMO}(S_{2\theta})}^{2\theta} + \|Z^1\|_{\text{BMO}(S_{2\theta})}^{2\theta} \right] \text{ a.s.}$$

so that $\|c\|_{\text{BMO}(S_2)} < \infty$. It remains to show that p_0 can be chosen such that $p_0 \in (1, 2)$. Here we repeat the above argument and check, for $0 \leq a < b \leq T$ and $\eta \in (\theta, 1)$, that

$$\begin{aligned} &\mathbb{E} \left(\int_a^b |c_s|^2 ds | \mathcal{F}_a \right) \\ &\leq L_Z^2 3^{2\theta} \left[(b - a) + (b - a)^{1 - \frac{\theta}{\eta}} [\|Z^0\|_{\text{BMO}(S_{2\eta})}^{2\theta} + \|Z^1\|_{\text{BMO}(S_{2\eta})}^{2\theta}] \right] \text{ a.s.} \end{aligned}$$

This yields $\lim_N \text{sl}_N^{S_2}(c) = 0$ and we can choose $p_0 \in (1, 2)$. Therefore we may apply Lemma 6.34 with $p = 2$ and this yields uniqueness.

The conclusion $s_\infty = 0$ follows by Remark 6.5 (2b), which is the same reasoning as used for $\lim_N \text{sl}_N^{S_2}(c) = 0$ above. \square

REMARK 6.14. Theorem 6.13 is an extension of the known case $\theta = 1$ (cf. [43, 57]). For $\theta = 1$ and $\xi \in L_\infty$ Theorem 6.6 gives a solution (Y, Z) with $\sup_{t \in [0, T]} \|Y_t\|_\infty < \infty$ and $|Z| \in \text{BMO}(S_2)$. Assuming two such solutions, we may follow the (second half of the) part about uniqueness in the proof of Theorem 6.13. Here the difference is that we only get *some* $p_0 \in (1, \infty)$ for applying Lemma 6.34. However, $|Z^0 - Z^1| \in \text{BMO}(S_2)$ implies that all moments of $\int_0^T |Z_s^0 - Z_s^1|^2 ds$ exist and Lemma 6.34 is applicable for any $p \in (p_0, \infty) \cap [2, \infty)$. Therefore, in the case $\theta = 1$ and $\xi \in L_\infty$ the solution (Y, Z) is unique when $\sup_{t \in [0, T]} \|Y_t\|_\infty < \infty$ and $|Z| \in \text{BMO}(S_2)$.

We finish by an example illustrating $\xi \in \text{cExp}$.

EXAMPLE 6.15. Let $d = 1$, $\eta \in (0, 1)$,

$$\varphi_\eta(t) := \log \left(1 + (T - t)^{1 - \frac{1}{\eta}} \right) \quad \text{for } t \in [0, T),$$

so that $\varphi_\eta(t) \uparrow \infty$ as $t \rightarrow T$ and define the stopping time

$$\tau_\eta := \inf \{t \in [0, T) : W_t = \varphi_\eta(t)\} \wedge T.$$

Let

$$e^\xi := 1 + e^{W_{\tau_\eta} - \frac{\tau_\eta}{2}}$$

so that $\xi(\omega) \in (0, \infty)$ and

$$\mathbb{E}(e^\xi | \mathcal{F}_t) = 1 + e^{W_{\tau_\eta \wedge t} - \frac{\tau_\eta \wedge t}{2}} \leq 2 + (T - t)^{1 - \frac{1}{\eta}} \text{ a.s.}$$

for $t \in [0, T)$. On the other hand, $\xi \notin L_\infty$ because for all $c > 0$ one has that $\mathbb{P}(W_{\tau_\eta} > c) > 0$. The latter fact can be checked by taking any $0 < \varepsilon < \varphi_\eta(0) < c < \infty$ and $S \in (0, T)$ with $c < \varphi_\eta(S)$ and using the known fact that $\mathbb{P}(\sup_{t \in [0, S]} |W_t| \leq \varepsilon) > 0$ so that the probability that the Brownian motion exceeds φ_η on $[S, (S + T)/2]$ is positive.

6.4. Settings for the stability theorem

The aim of this section is to discuss some settings for the stability Theorem 6.4.

6.4.1. Forward setting. This setting corresponds to the setting of stochastic integration. If the generator f does not depend on Y , then the process Y computes directly as

$$Y_t = Y_0 - \int_0^t f(s, Z_s) ds + \int_0^t Z_s dW_s.$$

This enables us to construct examples to understand what the correct conditions on Z in the quadratic case might be. Let us mention two cases:

- (a) Taking Z from Example 5.12 for $0 < \theta < \eta = 1$, we have examples where the Z -process fails to be in $\text{BMO}(S_2)$ but satisfies $Z \in \text{BMO}(S_{2\theta})$ and $\int_0^T |Z_s|^2 ds \in L_{\text{exp}}$. The latter enables us to apply Lemma 6.2 under suitable integrability conditions on $\int_0^T |f(s, 0)| ds$ (note that $L_{\text{exp}} \subseteq L_p$ for all $p \in (0, \infty)$).
- (b) Similarly, for $\theta = 1$ we obtain an L_p -solution of our BSDE under (B3), (B5), and $\left(\int_0^T |Z_t|^2 dt\right)^{\frac{1}{2}} \in L_{2p}$ (see the arguments at the end of Section 6.1). Therefore we can take any $Z \in \text{BMO}(S_2)$, in particular, Z can be an unbounded BMO-process in the quadratic setting.

6.4.2. Potential estimates for the generator. In applications of Theorem 6.4 one might need to estimate

$$\left\| \int_t^T |f^\varphi(s, Y_s^\psi, Z_s^\psi) - f^\psi(s, Y_s^\psi, Z_s^\psi)| ds \right\|_p$$

from above. One way to do this (we do not consider the remaining assumptions for Theorem 6.4) is to find a potential estimate

$$|f^\varphi(s, y, z) - f^\psi(s, y, z)| \leq |\langle (1, |y|, |z|, |z|^{1+\theta}), V_s^\varphi - V_s^\psi \rangle|$$

for all (s, y, z) where the potential $(V_s)_{s \in [0, T]}$ is a predictable process

$$V_s : \Omega \rightarrow \mathbb{R}^4.$$

Below we illustrate some special cases for V . The general construction is as follows: We consider a continuous

$$h : [0, T] \times \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R},$$

where $N \geq 1$, and a predictable \mathbb{R}^N -valued process $A = (A_t)_{t \in [0, T]}$ on Ω to let

$$f(t, \omega, y, z) := h(t, A_t(\omega), y, z).$$

Then f is $(\mathcal{P}, \mathcal{B}(C(\mathbb{R}^{1+d})))$ -measurable. Assume that $A^\varphi = (A_t^\varphi)_{t \in [0, T]}$ is a \mathcal{P}^φ -measurable transformation of \tilde{A} , where \tilde{A} is the canonical extension of A to $\bar{\Omega}$. We get that

$$f^\varphi(t, \bar{\omega}, y, z) := h(t, A_t^\varphi(\bar{\omega}), y, z)$$

is $(\mathcal{P}^\varphi, \mathcal{B}(C(\mathbb{R}^{1+d})))$ -measurable and, for any fixed $(y, z) \in \mathbb{R}^{1+d}$, that $f^\varphi(\cdot, \cdot, y, z) : [0, T] \times \bar{\Omega} \rightarrow \mathbb{R}$ is a transformation of the canonical extension $\tilde{f}(\cdot, \cdot, y, z) : [0, T] \times \bar{\Omega} \rightarrow \mathbb{R}$ (see Theorem 2.6(3) applied to $Y_{t,1} = t$ and $(Y_{t,2}(\bar{\omega}), \dots, Y_{t,N+1}(\bar{\omega})) = \tilde{A}_t(\bar{\omega})$). Therefore we will take in the sequel as transformed generator the map f^φ as defined above.

EXAMPLE 6.16. Let

$$f(s, \omega, y, z) := h(s, A_s(\omega), y, z),$$

where $h : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ is continuous with

$$|h(t, x_0, y_0, z_0) - h(t, x_1, y_1, z_1)| \leq L_X |x_0 - x_1| + L_Y |y_0 - y_1| + L_Z [1 + |z_0| + |z_1|] |z_0 - z_1|$$

for all $(t, x_0, x_1, y_0, y_1, z_0, z_1)$ and $(A_t)_{t \in [0, T]}$ is a predictable process. Then we get

$$|f^\varphi(s, y, z) - f^\psi(s, y, z)| \leq L_X |A_s^\varphi - A_s^\psi| \quad \text{and} \quad V_s := (L_X A_s, 0, 0, 0)$$

and

$$\left\| \int_t^T |f^\varphi(s, Y_s^\psi, Z_s^\psi) - f^\psi(s, Y_s^\psi, Z_s^\psi)| ds \right\|_p \leq L_X \left\| \int_t^T |A_s^\varphi - A_s^\psi| ds \right\|_p.$$

The next example indicates the case of random Lipschitz constants for y :

EXAMPLE 6.17. Assume that

$$f(s, \omega, y, z) := A_s(\omega)g(y)$$

where $g : \mathbb{R} \rightarrow \mathbb{R}$ is a Lipschitz function and $(A_s)_{s \in [0, T]}$ is predictable and uniformly bounded in (s, ω) . Then

$$|f^\varphi(s, y, z) - f^\psi(s, y, z)| \leq |g(y)| |A_s^\varphi - A_s^\psi| \leq [|g(0)| + \text{Lip}(g)|y|] |A_s^\varphi - A_s^\psi|$$

and $V_s := (|g(0)|A_s, \text{Lip}(g)A_s, 0, 0)$. Here we get (for example) that

$$\begin{aligned} & \left\| \int_t^T |f^\varphi(s, Y_s^\psi, Z_s^\psi) - f^\psi(s, Y_s^\psi, Z_s^\psi)| ds \right\|_p \\ & \leq |g(0)| \left\| \int_t^T |A_s^\varphi - A_s^\psi| ds \right\|_p + \text{Lip}(g) \left\| \int_t^T |Y_s^\psi| |A_s^\varphi - A_s^\psi| ds \right\|_p \end{aligned}$$

$$\begin{aligned}
&\leq |g(0)| \left\| \int_t^T |A_s^\varphi - A_s^\psi| ds \right\|_p + \text{Lip}(g) \left\| {}^*(Y^\psi)_t \int_t^T |A_s^\varphi - A_s^\psi| ds \right\|_p \\
&\leq |g(0)| \left\| \int_t^T |A_s^\varphi - A_s^\psi| ds \right\|_p + \text{Lip}(g) \| {}^*(Y^\psi)_t \|_{p_0} \left\| \int_t^T |A_s^\varphi - A_s^\psi| ds \right\|_{p_1} \\
&\leq [|g(0)| + \text{Lip}(g) \| {}^*Y_t \|_{p_0}] \left\| \int_t^T |A_s^\varphi - A_s^\psi| ds \right\|_{p_1}
\end{aligned}$$

for any $\frac{1}{p} = \frac{1}{p_0} + \frac{1}{p_1}$ with $p < p_0, p_1 < \infty$ and

$${}^*C_t := \sup_{s \in [t, T]} |C_s|,$$

where we used that ${}^*(Y^\psi)_t$ and *Y_t have the same distribution which follows from Theorem 2.6(2).

The last example concerns the Z component.

EXAMPLE 6.18. Assume that

$$f(s, \omega, y, z) := A_s(\omega) |z|^{1+\theta}$$

with $\theta \in (0, 1)$, where $(A_s)_{s \in [0, T]}$ is predictable and uniformly bounded in (s, ω) . Then

$$|f^\varphi(s, y, z) - f^\psi(s, y, z)| \leq |z|^{1+\theta} |A_s^\varphi - A_s^\psi|$$

and $V_s := (0, 0, 0, A_s)$. Because we have

$$\begin{aligned}
|f(s, \omega, y_0, z_0) - f(s, \omega, y_1, z_1)| &\leq |A_s(\omega)| \left| |z_0|^{1+\theta} - |z_1|^{1+\theta} \right| \\
&\leq [1 + \theta] |A_s(\omega)| \left| |z_0| - |z_1| \right| [1 + |z_0| + |z_1|]^\theta \\
&\leq [1 + \theta] |A_s(\omega)| |z_0 - z_1| [1 + |z_0| + |z_1|]^\theta
\end{aligned}$$

condition (B3) is satisfied. Then an upper bound is obtained by

$$\begin{aligned}
&\left\| \int_t^T |f^\varphi(s, Y_s^\psi, Z_s^\psi) - f^\psi(s, Y_s^\psi, Z_s^\psi)| ds \right\|_p \\
&\leq \left\| \int_t^T |Z_s^\psi|^{1+\theta} |A_s^\varphi - A_s^\psi| ds \right\|_p \\
&\leq \left\| \left(\int_t^T |Z_s^\psi|^2 ds \right)^{\frac{1+\theta}{2}} \left(\int_t^T |A_s^\varphi - A_s^\psi|^{\frac{2}{1-\theta}} ds \right)^{\frac{1-\theta}{2}} \right\|_p \\
&\leq \left\| \left(\int_t^T |Z_s^\psi|^2 ds \right)^{\frac{1}{2}} \right\|_{(1+\theta)p_0}^{1+\theta} \left\| \left(\int_t^T |A_s^\varphi - A_s^\psi|^{\frac{2}{1-\theta}} ds \right)^{\frac{1-\theta}{2}} \right\|_{p_1}
\end{aligned}$$

for any $\frac{1}{p} = \frac{1}{p_0} + \frac{1}{p_1}$ with $p < p_0, p_1 < \infty$, where we use Remark 2.7(2) in the last step.

6.4.3. Theorem 6.4 for the perturbation $(\varphi, \psi) = (\chi_{(a,b]}, 0)$. The importance of the pair $(\varphi, \psi) = (\chi_{(a,b]}, 0)$ follows from the fact that

$$\|Y_t - Y_t^{(t-\varepsilon, t]}\|_p \sim_2 \|Y_t - \mathbb{E}(Y_t | \mathcal{F}_{t-\varepsilon})\|_p$$

for $p \in [1, \infty]$, i.e. the fractional smoothness of Y_t is measured in terms of the speed of convergence of the conditional expectations. In the case $(\varphi, \psi) = (\chi_{(a,b]}, 0)$ we have that (27) implies two inequalities that give different information about the L_p -variation of the processes $Y = (Y_t)_{t \in [0, T]}$ and $Z = (Z_t)_{t \in [0, T]}$: Firstly, for $0 < \varepsilon < t$ we have that

$$(40) \quad \left\| Y_t - Y_t^{(t-\varepsilon, t]} \right\|_p + \left\| \left(\int_t^T |Z_s - Z_s^{(t-\varepsilon, t]}|^2 ds \right)^{\frac{1}{2}} \right\|_p \\ \leq c_{(6.4)} \left[\left\| \xi - \xi^{(t-\varepsilon, t]} \right\|_p + \left\| \int_t^T |f(s, Y_s, Z_s) - f^{(t-\varepsilon, t]}(s, Y_s, Z_s)| ds \right\|_p \right]$$

and, secondly,

$$(41) \quad \left\| \left(\int_{t-\varepsilon}^t |Z_s|^2 ds \right)^{\frac{1}{2}} \right\|_p \\ \leq c_{(6.4)} \left[\left\| \xi - \xi^{(t-\varepsilon, t]} \right\|_p + \left\| \int_{t-\varepsilon}^T |f(s, Y_s, Z_s) - f^{(t-\varepsilon, t]}(s, Y_s, Z_s)| ds \right\|_p \right].$$

6.4.4. Theorem 6.4 and Besov spaces. We want to transform Theorem 6.4 into an embedding theorem for the Besov spaces \mathbb{B}_p^Φ . As the BSDEs we consider might be even quadratic we have - in some sense - a non-linear embedding theorem. To handle the assumption on the generator we need a slight extension of our anisotropic Besov spaces:

DEFINITION 6.19. For $q, r \in [1, \infty)$, a predictable process $(A_t)_{t \in [0, T]}$ with

$$\left\| \left(\int_0^T |A_s|^r ds \right)^{\frac{1}{r}} \right\|_q < \infty,$$

for $t \in [0, T]$, and for an admissible functional Φ we let

$$\|A\|_{\Phi, q}^{r, t} := \Phi \left(\psi \rightarrow \left\| \left(\int_t^T |A_s - A_s^\psi|^r ds \right)^{\frac{1}{r}} \right\|_q \right).$$

First we show that this definition is possible:

LEMMA 6.20. *The map*

$$\psi \rightarrow \left\| \left(\int_t^T |A_s - A_s^\psi|^r ds \right)^{\frac{1}{r}} \right\|_q$$

is continuous as a map from Δ into $[0, \infty)$.

PROOF. We fix an $N \geq 1$ and consider the truncation $A_t^N := (-N) \vee (A_t \wedge N)$. For $u := q \vee r$ and $\psi_n, \psi \in \Delta$ we get

$$\begin{aligned}
& \left\| \left(\int_t^T |A_s^{\psi_n} - A_s^\psi|^r ds \right)^{\frac{1}{r}} \right\|_q \\
& \leq \left\| \left(\int_t^T |A_s^{\psi_n} - (A^N)_s^{\psi_n}|^r ds \right)^{\frac{1}{r}} \right\|_q \\
& \quad + \left\| \left(\int_t^T |(A^N)_s^{\psi_n} - (A^N)_s^\psi|^r ds \right)^{\frac{1}{r}} \right\|_q \\
& \quad + \left\| \left(\int_t^T |(A^N)_s^\psi - A_s^\psi|^r ds \right)^{\frac{1}{r}} \right\|_q \\
& = 2 \left\| \left(\int_t^T |A_s - (A^N)_s|^r ds \right)^{\frac{1}{r}} \right\|_q + \left\| \left(\int_t^T |(A^N)_s^{\psi_n} - (A^N)_s^\psi|^r ds \right)^{\frac{1}{r}} \right\|_q \\
& \leq 2 \left\| \left(\int_t^T |A_s - (A^N)_s|^r ds \right)^{\frac{1}{r}} \right\|_q + c_{q,r,T} \left(\int_t^T \|(A^N)_s^{\psi_n} - (A^N)_s^\psi\|_u^u ds \right)^{\frac{1}{u}}
\end{aligned}$$

where we used for the equality Remark 2.7(2). Applying dominated convergence twice we get that

$$\lim_N \left\| \left(\int_t^T |A_s - (A^N)_s|^r ds \right)^{\frac{1}{r}} \right\|_q = 0.$$

Moreover, using Theorem 2.6(6) we find a Borel set $B \subseteq [0, T]$ of Lebesgue measure T such that A_t^ρ is the transformation of A_t for any $t \in B$ and $\rho \in \{\psi, \psi_1, \psi_2, \dots\}$. In case $\psi_n \rightarrow \psi$ we can therefore apply Lemma 4.7 to conclude the proof because this implies that

$$\lim_n \left(\int_t^T \|(A^N)_s^{\psi_n} - (A^N)_s^\psi\|_u^u ds \right)^{\frac{1}{u}} = 0. \quad \square$$

Now we obtain the following embedding theorem:

COROLLARY 6.21. *Assume that the assumptions of Theorem 6.4 are satisfied, $t \in [0, T]$, and that there are predictable processes $(V_s^l)_{s \in [t, T]}$ such that, for all $\psi \in \Delta$,*

$$\left\| \int_t^T |f(s, Y_s^\psi, Z_s^\psi) - f^\psi(s, Y_s^\psi, Z_s^\psi)| ds \right\|_p \leq \sum_{l=1}^L \|V^l - (V^l)^\psi\|_{L_{q_l}(L_{r_l}([t, T]))}$$

for some $q_l \in [p, \infty)$ and $r_l \in [1, \infty)$ ¹. Let $\Phi : C^+(\Delta) \rightarrow [0, \infty]$ be admissible in the sense of Definition 4.10. Then we have that

$$\|Y_t\|_{\Phi,p} + \|Z\|_{\Phi,p}^{2,t} \leq 2c_{(6.4)} \left[\|\xi\|_{\Phi,p} + \sum_{l=1}^L \|V^l\|_{\Phi,q_l}^{r_l,t} \right].$$

PROOF. The statement follows directly from Theorem 6.4 applied to the pair $(0, \psi)$. \square

Examples, how to obtain processes $(V_s^l)_{s \in [t, T]}$, can be found in Section 6.4.2. For the sake of illustration we first combine Corollary 6.21 with Theorem 6.13 (note that we use conditions (B3) and (B8)) so that the assumptions of Theorem 6.4 are automatically satisfied with $p = 2$ and $s_\infty = 0$:

COROLLARY 6.22. Assume that $\theta \in (0, 1)$, $t \in [0, T]$, $\xi \in \text{cExp}$, and that (Y, Z) is the unique solution to the BSDE (23) obtained in Theorem 6.13. Suppose a predictable process $(V_s)_{s \in [t, T]}$ such that

$$\left\| \int_t^T \sup_{y,z} |f(s, y, z) - f^\psi(s, y, z)| ds \right\|_2 \leq \|V - (V)^\psi\|_{L_2(L_1([t, T]))}$$

for all $\psi \in \Delta$. Let $\Phi : C^+(\Delta) \rightarrow [0, \infty]$ be admissible in the sense of Definition 4.10. Then we have that

$$\|Y_t\|_{\Phi,2} + \|Z\|_{\Phi,2}^{2,t} \leq 2c_{(6.4)} \left[\|\xi\|_{\Phi,2} + \|V\|_{\Phi,2}^{1,t} \right].$$

Taking also Theorem 4.19 into the account we obtain another version of Corollary 6.22 that only uses that ξ is locally in $\mathbb{D}_{1,2}$ in the sense to check perturbations of the Gaussian structure up to time t only. This confirms the smoothing effect of a BSDE as this already implies the smoothness of Y_t . More precisely we get:

COROLLARY 6.23. Assume that $\theta \in (0, 1)$, $t \in [0, T]$, $\xi \in \text{cExp}$, and that (Y, Z) is the unique solution to the BSDE (23) obtained in Theorem 6.13. Then we have

$$(42) \quad \text{esssup}_{s \in [0, t]} \|D_s Y_t\|_2 \leq c \sup_{0 \leq a < b \leq t} \frac{1}{\sqrt{b-a}} \left[\|\xi - \xi^{(a,b)}\|_2 + \left\| \int_t^T \sup_{y,z} |f(s, y, z) - f^{(a,b)}(s, y, z)| ds \right\|_2 \right]$$

with $c := c_{(6.4)} c_{(4.19)(1),2}$ and $c_{(4.19)(1),2} \geq 1$ taken from Theorem 4.19 in the sense that if the right-hand side is finite, then $Y_t \in \mathbb{D}_{1,2}$ and (42) holds.

PROOF. We apply Theorems 6.4 and 4.19, where for the latter we use

$$Y_t^{(a,b)} = Y_t^{(a \wedge t, b \wedge t)} \text{ a.s.}$$

because Y_t is \mathcal{F}_t -measurable. \square

¹The V^l may depend on $(\xi, f, Y, Z, p, q_l, r_l)$.

If $\xi \in \mathbb{D}_{1,2}$, then we have that

$$\sup_{0 \leq a < b \leq t} \frac{\|\xi - \xi^{(a,b)}\|_2}{\sqrt{b-a}} \leq 2c_{(A.7)} \text{esssup}_{s \in [0,t]} \|D_s \xi\|_2$$

by Corollary 6.29, but for Corollary 6.23 the assumption $\xi \in \mathbb{D}_{1,2}$ is not necessary.

6.5. On the L_p -variation of BSDEs

In this section we show how Theorem 6.4 can be applied in order to obtain information about the L_p -variation of our BSDE. The link between the L_p -variation of the Y -process and the stability result Theorem 6.4 consists in the observation

$$\|A_t - A_s\|_p \leq \|A_t - \mathbb{E}(A_t | \mathcal{F}_s)\|_p + \|\mathbb{E}(A_t | \mathcal{F}_s) - A_s\|_p \leq 3\|A_t - A_s\|_p,$$

where $p \in [1, \infty]$, $(A_t)_{t \in [0,T]} \subseteq \mathcal{L}_p$ is adapted, and $0 \leq s \leq t \leq T$. Our estimate for the Z -process will follow directly from Theorem 6.4.

In Remark 6.33(1) below we show that under the conditions $\int_0^T \|Z_r\|_p^2 dr < \infty$ and $\int_0^T \|f(r, Y_r, Z_r)\|_p dr < \infty$, and under the *a-priori* knowledge of the behaviour of the functions $r \rightarrow \|Z_r\|_p$ and $r \rightarrow \|f(r, Y_r, Z_r)\|_p$ one gets a rate of $1/\sqrt{n}$ for the L_p -variation of Y and Z by adapted time-nets. In Theorem 6.32 below we will deduce estimates with explicit adapted time-nets where we only assume conditions on the initial data (ξ, f) . Regarding the case $p \in (2, \infty)$ there is another aspect: In Remark 6.33(2) we show that even for the zero generator case one might have situations where one cannot achieve the rate $1/\sqrt{n}$ for the variation of Y , i.e. the variation of Y is asymptotically higher. Our sufficient conditions give cases where one gets the rate $1/\sqrt{n}$ for the case $p \in (2, \infty)$.

In the following the random variables are considered on the product space $\overline{\Omega}$ if necessary. In particular, random variables defined on Ω are automatically extended to $\overline{\Omega}$ in the natural way when needed.

THEOREM 6.24. *Suppose that the assumptions of Theorem 6.4 are satisfied. Then, for $c_{(6.24)} := c_{(6.4)}[1 + c_{(5.21),p}L_Z(\sqrt{T} + 1)]$ and $0 \leq s < t \leq T$, one has*

$$\begin{aligned} & \left\| \sup_{r \in [s,t]} |Y_r - Y_s| \right\|_p + \left\| \left(\int_s^t |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ & \leq \left\| \int_s^t |f(r, 0, 0)| dr \right\|_p + L_Y(t-s) \sup_{r \in [0,T]} \|Y_r\|_p + c_{(6.24)} [1 + \|Z\|_{\text{BMO}(S_2)}^\theta] \times \\ & \quad \times \left[\|\xi - \xi^{(s,t)}\|_p + \left\| \int_s^T |f(r, Y_r, Z_r) - f^{(s,t)}(r, Y_r, Z_r)| dr \right\|_p \right]. \end{aligned}$$

PROOF. We fix $0 \leq s < t \leq T$ and remark that $\left\| \int_s^t |f(r, 0, 0)| dr \right\|_p < \infty$ according to condition (B5). We let $(q_l)_{l=1}^\infty$ be an enumeration of the rational numbers from $(s, t]$ so that

$$\left\| \sup_{r \in [s,t]} |Y_r - Y_s| \right\|_p = \sup_{m=1,2,\dots} \left\| \sup_{q \in \{q_1, \dots, q_m\}} |Y_q - Y_s| \right\|_p$$

by monotone convergence. Using Lemma 4.20, the fact that the Y_{q_i} are \mathcal{F}_t -measurable, and Theorem 6.4 we obtain that

$$\begin{aligned}
& \left\| \sup_{q \in \{q_1, \dots, q_m\}} |Y_q - Y_s| \right\|_p + \left\| \left(\int_s^t |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\
& \leq \left\| \sup_{q \in \{q_1, \dots, q_m\}} |Y_q - \mathbb{E}^{\mathcal{F}_s} Y_q| \right\|_p + \left\| \sup_{q \in \{q_1, \dots, q_m\}} |\mathbb{E}^{\mathcal{F}_s} Y_q - Y_s| \right\|_p \\
& \quad + \left\| \left(\int_s^t |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\
& \leq \left\| \sup_{q \in \{q_1, \dots, q_m\}} |Y_q - Y_q^{(s,t)}| \right\|_p + \left\| \left(\int_s^t |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\
& \quad + \left\| \sup_{q \in \{q_1, \dots, q_m\}} |\mathbb{E}^{\mathcal{F}_s} Y_q - Y_s| \right\|_p \\
& \leq c_{(6.4)} \left[\|\xi - \xi^{(s,t)}\|_p + \left\| \int_s^T |f(r, Y_r, Z_r) - f^{(s,t)}(r, Y_r, Z_r)| dr \right\|_p \right] \\
& \quad + \left\| \sup_{q \in \{q_1, \dots, q_m\}} |\mathbb{E}^{\mathcal{F}_s} Y_q - Y_s| \right\|_p.
\end{aligned}$$

By Corollary 5.21 and (41) we bound the last term by

$$\begin{aligned}
& \left\| \sup_{q \in \{q_1, \dots, q_m\}} |\mathbb{E}^{\mathcal{F}_s} Y_q - Y_s| \right\|_p \\
& = \left\| \sup_{q \in \{q_1, \dots, q_m\}} \left| \mathbb{E}^{\mathcal{F}_s} \int_s^q f(r, Y_r, Z_r) dr \right| \right\|_p \\
& \leq \left\| \int_s^t |f(r, Y_r, Z_r)| dr \right\|_p \\
& \leq \left\| \int_s^t |f(r, 0, 0)| dr \right\|_p + L_Y \left\| \int_s^t |Y_r| dr \right\|_p + L_Z \left\| \int_s^t [1 + |Z_r|]^\theta |Z_r| dr \right\|_p \\
& \leq \left\| \int_s^t |f(r, 0, 0)| dr \right\|_p + L_Y(t-s) \sup_{r \in [s,t]} \|Y_r\|_p \\
& \quad + L_Z c_{(5.21),p} \left\| \left(\int_s^t |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \left\| (\chi_{(s,t]}(r) [1 + |Z_r|]^\theta)_{r \in [0,T]} \right\|_{\text{BMO}(S_2)} \\
& \leq \left\| \int_s^t |f(r, 0, 0)| dr \right\|_p + L_Y(t-s) \sup_{r \in [s,t]} \|Y_r\|_p \\
& \quad + L_Z c_{(5.21),p} c_{(6.4)} \left[\|\xi - \xi^{(s,t)}\|_p + \left\| \int_s^T |f(r, Y_r, Z_r) - f^{(s,t)}(r, Y_r, Z_r)| dr \right\|_p \right]
\end{aligned}$$

$$\times [\sqrt{t-s} + \| |Z|^\theta \|_{\text{BMO}(S_2)}].$$

As remarked in the beginning of the proof of Theorem 6.4 we have $\sup_{r \in [0, T]} |Y_r| \in \mathcal{L}_p$ so that $\alpha_p := L_Y \sup_{r \in [0, T]} \|Y_r\|_p < \infty$. Therefore,

$$\begin{aligned} \left\| \sup_{q \in \{q_1, \dots, q_m\}} |\mathbb{E}^{\mathcal{F}_s} Y_q - Y_s| \right\|_p &\leq \left\| \int_s^t |f(r, 0, 0)| dr \right\|_p + \alpha_p(t-s) \\ &\quad + \beta_p \left[\|\xi - \xi^{(s, t]}\|_p + \left\| \int_s^T |f(r, Y_r, Z_r) - f^{(s, t]}(r, Y_r, Z_r)| dr \right\|_p \right] \end{aligned}$$

for $\beta_p := L_Z c(5.21), p c(6.4) [\sqrt{T} + \| |Z|^\theta \|_{\text{BMO}(S_2)}]$. \square

The variation of our BSDE we measure by the following quantity:

DEFINITION 6.25. Let $p \in [1, \infty)$, $A = (A_t)_{t \in [0, T]}$ be a measurable càdlàg process $A : [0, T] \times \Omega \rightarrow \mathbb{R}$, and $C = (C_t)_{t \in [0, T]}$ be a measurable process $C : [0, T] \times \Omega \rightarrow \mathbb{R}^d$, where \mathbb{R}^d is equipped with the euclidean norm. For a deterministic time-net $\tau = (t_i)_{i=0}^n$ with $0 = t_0 \leq t_1 \leq \dots \leq t_n = T$ we let

$$\text{var}_p([A, C]|\tau) := \sup_{i=1, \dots, n} \left\| \sup_{t_{i-1} \leq s \leq t \leq t_i} |A_t - A_s| \right\|_p + \sup_{i=1, \dots, n} \left\| \left(\int_{t_{i-1}}^{t_i} |C_r|^2 dr \right)^{\frac{1}{2}} \right\|_p.$$

The variation $\text{var}_p([A, C]|\tau)$ behaves sub-additive as expected:

LEMMA 6.26. For $p \in [1, \infty)$, families $((A_t^j, C_t^j))_{t \in [0, T]}$ and time-nets τ^j , $j = 0, 1$, as in Definition 6.25, one has that

$$\text{var}_p([A^0 + A^1, C^0 + C^1]|\tau^0 \cup \tau^1) \leq \text{var}_p([A^0, C^0]|\tau^0) + \text{var}_p([A^1, C^1]|\tau^1).$$

PROOF. Assume that $\tau = (t_i)_{i=0}^{n_0+n_1-1}$ is an ordering of the union of $\tau^0 = (t_i^0)_{i=0}^{n_0}$ and $\tau^1 = (t_i^1)_{i=0}^{n_1}$. Then one has that the interval $[t_{i-1}, t_i]$ is contained in a closed interval of τ^0 and, at the same time, in a closed interval of τ^1 , so that

$$\begin{aligned} &\left\| \sup_{t_{i-1} \leq s \leq t \leq t_i} |(A_t^0 + A_t^1) - (A_s^0 + A_s^1)| \right\|_p + \left\| \left(\int_{t_{i-1}}^{t_i} |C_r^0 + C_r^1|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ &\leq \left\| \sup_{t_{i-1} \leq s \leq t \leq t_i} |A_t^0 - A_s^0| \right\|_p + \left\| \left(\int_{t_{i-1}}^{t_i} |C_r^0|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ &\quad + \left\| \sup_{t_{i-1} \leq s \leq t \leq t_i} |A_t^1 - A_s^1| \right\|_p + \left\| \left(\int_{t_{i-1}}^{t_i} |C_r^1|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ &\leq \text{var}_p([A^0, C^0]|\tau^0) + \text{var}_p([A^1, C^1]|\tau^1). \end{aligned}$$

\square

Now we formulate consequences of Theorem 6.24 in two different scenarios: The first Theorem 6.31 still relies on the assumptions of Theorem 6.4. In the next step

Theorem 6.31 will be combined with the results from Section 6.3 to guarantee the validity of the assumptions of Theorem 6.4. This yields to Theorem 6.32.

To shorten the formulation of the statements we work with the following two definitions.

The first definition extends the spaces \mathbb{B}_p^Φ to the initial data (ξ, f) of the BSDE:

DEFINITION 6.27. We say that $(\xi, f) \in \mathbb{B}_p^{\Phi_{\gamma, \Gamma}}$, where $p \in (0, \infty)$, $\gamma \in [2, \infty)$, and $\Gamma : [0, T] \rightarrow [0, \infty)$ is integrable, provided that $\xi \in L_p$ and for all $0 \leq a < b \leq T$,

$$\|\xi - \xi^{(a,b)}\|_p + \left\| \int_a^T \sup_{(y,z) \in \mathbb{R}^{d+1}} |f(r, y, z) - f^{(a,b)}(r, y, z)| dr \right\|_p \leq \left(\int_a^b \Gamma(r) dr \right)^{\frac{1}{\gamma}}.$$

The term $\int_a^T \sup_{(y,z) \in \mathbb{R}^{d+1}} |f(r, y, z) - f^{(a,b)}(r, y, z)| dr$ is an extended random variable on $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$. Concerning the generator, the above definition reflects the situation described in Example 6.16, where the generator f is obtained from some appropriate h with

$$f(r, \omega, y, z) := h(r, A_r(\omega), y, z).$$

The second definition recalls a well know principle to generate adapted time-nets:

DEFINITION 6.28. Letting $\Lambda : [0, T] \rightarrow (0, \infty)$ be integrable and $n \geq 1$, the time net τ_n^Λ consists of $0 = t_0 < \dots < t_n = T$ such that, for all $i = 1, \dots, n$,

$$\int_{t_{i-1}}^{t_i} \Lambda(r) dr = \frac{1}{n} \int_0^T \Lambda(r) dr.$$

The following corollary, which follows directly from Lemma 4.20 and (15), yields to the fundamental example for $\gamma = 2$ concerning the part $\|\xi - \xi^{(a,b)}\|_p$ in Definition 6.27 above:

COROLLARY 6.29. For $p \in [2, \infty)$ and $\xi \in \mathbb{D}_{1,2} \cap L_p$ with $\int_{(0,T]} \|D_r \xi\|_p^2 dr < \infty$ one has for all $0 \leq a < b \leq T$ that

$$\|\xi - \xi^{(a,b)}\|_p \leq 2c_{(A.7)} \left(\int_{(a,b]} \|D_r \xi\|_p^2 dr \right)^{\frac{1}{2}}.$$

REMARK 6.30. For $\Gamma > 0$ Definition 6.27 and Example 4.11 yield to the admissible functional

$$\Phi_{\gamma, \Gamma}(F) := \sup_{0 \leq a < b \leq T} \frac{F(\chi_{(a,b]})}{\sqrt[\gamma]{\int_a^b \Gamma(r) dr}}$$

that recovers the functional Φ_γ from (16) by $\Gamma \equiv 1$.

Our first corollary of Theorem 6.24 is

THEOREM 6.31. *Let $p, \gamma \in [2, \infty)$ and $\Gamma : [0, T] \rightarrow [0, \infty)$ be integrable. Suppose the assumptions of Theorem 6.4, $(\xi, f) \in \mathbb{B}_p^{\Phi, \gamma, \Gamma}$ and $\int_0^T \|f(r, 0, 0)\|_p dr < \infty$. Then*

$$\text{var}_p([Y, Z]|\tau_n^\Lambda) \leq \frac{c_{(6.31)}}{n} + \frac{d_{(6.31)}}{\sqrt[n]{n}}$$

for $\Lambda(r) := 1 + \|f(r, 0, 0)\|_p + \Gamma(r)$ and

$$c_{(6.31)} := 2\|\Lambda\|_{L_1([0, T])} [1 + L_Y \sup_{t \in [0, T]} \|Y_t\|_p],$$

$$d_{(6.31)} := 2c_{(6.24)} \|\Lambda\|_{L_1([0, T])}^{\frac{1}{\gamma}} [1 + \| |Z|^\theta \|_{\text{BMO}(S_2)}].$$

PROOF. For $0 \leq s < t \leq T$ Theorem 6.24 implies that

$$\begin{aligned} & \left\| \sup_{r \in [s, t]} |Y_r - Y_s| \right\|_p + \left\| \left(\int_s^t |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ & \leq \left\| \int_s^t |f(r, 0, 0)| dr \right\|_p + L_Y(t-s) \sup_{r \in [0, T]} \|Y_r\|_p \\ & \quad + c_{(6.24)} [1 + \| |Z|^\theta \|_{\text{BMO}(S_2)}] \times \\ & \quad \times \left[\|\xi - \xi^{(s, t]}\|_p + \left\| \int_s^T |f(r, Y_r, Z_r) - f^{(s, t]}(r, Y_r, Z_r)| dr \right\|_p \right] \\ & \leq \int_s^t \|f(r, 0, 0)\|_p dr + L_Y(t-s) \sup_{r \in [0, T]} \|Y_r\|_p \\ & \quad + c_{(6.24)} [1 + \| |Z|^\theta \|_{\text{BMO}(S_2)}] \left(\int_s^t \Gamma(r) dr \right)^{\frac{1}{\gamma}}. \end{aligned}$$

Assuming $0 = t_0 \leq t_1 \leq \dots \leq t_n = T$ we conclude by

$$\begin{aligned} & \left\| \sup_{t_{i-1} \leq s \leq t \leq t_i} |Y_t - Y_s| \right\|_p + \left\| \left(\int_{t_{i-1}}^{t_i} |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ & \leq 2 \left\| \sup_{r \in [t_{i-1}, t_i]} |Y_r - Y_{t_{i-1}}| \right\|_p + \left\| \left(\int_{t_{i-1}}^{t_i} |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ & \leq 2 \left[\int_{t_{i-1}}^{t_i} \|f(r, 0, 0)\|_p dr + L_Y(t_i - t_{i-1}) \sup_{r \in [0, T]} \|Y_r\|_p \right. \\ & \quad \left. + c_{(6.24)} [1 + \| |Z|^\theta \|_{\text{BMO}(S_2)}] \left(\int_{t_{i-1}}^{t_i} \Gamma(r) dr \right)^{\frac{1}{\gamma}} \right]. \end{aligned}$$

□

THEOREM 6.32. *Assume $\gamma \in [2, \infty)$ and that one of the following sets of conditions is satisfied:*

- (1) $\theta = 0$, $p \in [2, \infty)$, $\xi \in L_p$, (B3), $\int_0^T \|f(r, 0, 0)\|_p dr < \infty$, $(\xi, f) \in \mathbb{B}_p^{\Phi, \gamma, \Gamma}$.
- (2) $\theta \in (0, 1)$, $\xi \in \text{cExp}$, (B3), (B8), and $(\xi, f) \in \mathbb{B}_2^{\Phi, \gamma, \Gamma}$.
- (3) $\theta = 1$, $\xi \in L_\infty$, (B3), (B8), and $(\xi, f) \in \bigcap_{q \in [2, \infty)} \mathbb{B}_q^{\Phi, \gamma, \Gamma}$.

Define the weight function

$$\Lambda(r) := 1 + \|f(r, 0, 0)\|_u + \Gamma(r)$$

where $u = p$ for $\theta = 0$, $u = 2$ for $\theta \in (0, 1)$, and $u = \infty$ for $\theta = 1$. Then one has that

$$\sup_{n \geq 1} \sqrt[n]{n} \text{var}_v([Y, Z] | \tau_n^\Lambda) < \infty$$

for $v = p$ if $\theta = 0$, $v = 2$ if $\theta \in (0, 1)$, and for all $v \in (0, \infty)$ if $\theta = 1$, where for $\theta = 0$ the solution is taken from [18, Theorem 4.2], for $\theta \in (0, 1)$ from Theorem 6.13, and for $\theta = 1$ from Remark 6.14.

PROOF. The statement follows by a combination of Table 1 (cases I, II, and III) and Theorem 6.31. For part (3) we remark that we first deduce our statement for $v \in [2, \infty) \cap (p_0, \infty)$ with p_0 taken from Theorem 6.4, and then (obviously) the conclusion follows for all $v \in (0, \infty)$. \square

REMARK 6.33.

- (1) For $p \in [2, \infty)$ assume for our BSDE the conditions $\int_0^T \|Z_r\|_p^2 dr < \infty$ and $\int_0^T \|f(r, Y_r, Z_r)\|_p dr < \infty$. Take a net $\tau^n = (t_i^n)_{i=1}^n$ that satisfies

$$\int_{t_{i-1}^n}^{t_i^n} [\|f(r, Y_r, Z_r)\|_p + \|Z_r\|_p^2] dr = \frac{1}{n} \int_0^T [\|f(r, Y_r, Z_r)\|_p + \|Z_r\|_p^2] dr.$$

Given $i \in \{1, \dots, n\}$ we derive

$$\begin{aligned} & \left\| \sup_{t_{i-1}^n \leq s \leq t \leq t_i^n} |Y_t - Y_s| \right\|_p + \left\| \left(\int_{t_{i-1}^n}^{t_i^n} |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ & \leq \left\| \int_{t_{i-1}^n}^{t_i^n} |f(r, Y_r, Z_r)| dr \right\|_p + \left\| \sup_{q \in [t_{i-1}^n, t_i^n]} \left| \int_{t_{i-1}^n}^q Z_r dW_r \right| \right\|_p \\ & \quad + \left\| \left(\int_{t_{i-1}^n}^{t_i^n} |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ & \leq \int_{t_{i-1}^n}^{t_i^n} \|f(r, Y_r, Z_r)\|_p dr + [2\beta_p + 1] \left\| \left(\int_{t_{i-1}^n}^{t_i^n} |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ & \leq \int_{t_{i-1}^n}^{t_i^n} \|f(r, Y_r, Z_r)\|_p dr + [2\beta_p + 1] \left(\int_{t_{i-1}^n}^{t_i^n} \|Z_r\|_p^2 dr \right)^{\frac{1}{2}} \\ & \leq \frac{1}{n} \int_0^T [\|f(r, Y_r, Z_r)\|_p + \|Z_r\|_p^2] dr \\ & \quad + \frac{2\beta_p + 1}{\sqrt{n}} \left(\int_0^T [\|f(r, Y_r, Z_r)\|_p + \|Z_r\|_p^2] dr \right)^{\frac{1}{2}} \end{aligned}$$

where the Burkholder-Davis-Gundy inequalities (8) were exploited. Therefore we have a variation of $1/\sqrt{n}$ by taking the nets τ^n .

- (2) However, in general for $p \in (2, \infty)$ such an estimate is not always possible as shown by the following example for $d = 1$: Take an infinite time net converging to T ,

$$0 = t_0 < t_1 < t_2 < \dots,$$

and pair-wise disjoint $A_k \in \mathcal{F}_{t_k}$ of positive measure for $k = 1, 2, \dots$ (Given pair-wise disjoint non-empty finite intervals $I_k = (a_k, b_k)$ one can choose $A_1 := \{W_{t_1} \in I_1\}$ and $A_k := \{W_{t_1} \notin I_1, \dots, W_{t_{k-1}} \notin I_{k-1}, W_{t_k} \in I_k\}$ for $k \geq 2$.) For $(\alpha_k)_{k=1}^\infty \subset (0, \infty)$ and $s \in [0, T]$ define

$$\lambda_s := \sum_{k=2}^{\infty} \alpha_{k-1} \chi_{A_{k-1}} \chi_{(t_{k-1}, t_k]}(s).$$

Let $0 < \alpha < \frac{p}{2} - 1$ and arrange the α_k such that

$$\left\| \int_{(t_{k-1}, t_k]} \lambda_s dW_s \right\|_p = k^{-\frac{1+\alpha}{p}}$$

which implies

$$\mathbb{E} \left| \int_{(0, T]} \lambda_s dW_s \right|^p = \sum_{k=2}^{\infty} k^{-(1+\alpha)} < \infty.$$

Let us assume $c > 0$ and a sequence of time-nets τ^n , $0 = t_0^n \leq \dots \leq t_n^n = T$, such that

$$\|Y_{t_i^n} - Y_{t_{i-1}^n}\|_p \leq \frac{c}{\sqrt{n}} \quad \text{for } Y_t := \int_0^t \lambda_s dW_s.$$

Then $(t_{k-1}, t_k) \cap \tau^n = \emptyset$ for $k \geq 2$ implies that

$$\|Y_{t_k} - Y_{t_{k-1}}\|_p = k^{-\frac{1+\alpha}{p}} \leq \frac{c}{\sqrt{n}}$$

or, equivalently, the condition $k^{-\frac{1+\alpha}{p}} > \frac{c}{\sqrt{n}}$ gives $(t_{k-1}, t_k) \cap \tau^n \neq \emptyset$ for $k \geq 2$. In other words, all intervals (t_{k-1}, t_k) with

$$2 \leq k < \left(\frac{\sqrt{n}}{c} \right)^{\frac{p}{1+\alpha}}$$

contain at least one element of the time-net τ^n . This gives a contradiction to $\frac{p}{2(1+\alpha)} > 1$.

6.6. An a-priori estimate for BSDEs

In this section we follow the ideas of [19, Proof of Proposition 2.3] but adapt and extend the ideas for our purpose. Let $B = (B_t)_{t \in [0, T]}$ be an n -dimensional standard Brownian motion (where all paths are continuous) on a basis $(A, \mathcal{A}, \mathbb{Q}, (\mathcal{A}_t)_{t \in [0, T]})$, where $(A, \mathcal{A}, \mathbb{Q})$ is complete, $(\mathcal{A}_t)_{t \in [0, T]}$ the augmentation of the natural filtration of B , and $\mathcal{A}_T = \mathcal{A}$. We consider the two backward equations

$$\begin{aligned} Y_t^0 &= \xi^0 + \int_t^T f^0(s, Y_s^0, Z_s^0) ds - \int_t^T Z_s^0 dB_s, \\ Y_t^1 &= \xi^1 + \int_t^T f^1(s) ds - \int_t^T Z_s^1 dB_s, \end{aligned}$$

where we assume the following conditions:

- (D1) The processes f^1 , Z^0 and Z^1 are predictable and the processes Y^0 and Y^1 continuous and adapted,
- (D2) $\mathbb{E}|\xi^i|^2 < \infty$ and $\mathbb{E} \int_0^T |Z_s^i|^2 ds < \infty$ for $i = 0, 1$,
- (D3) $\mathbb{E} \left| \int_0^T |f^0(s, Y_s^0, Z_s^0)| ds \right|^2 < \infty$ and $\mathbb{E} \left| \int_0^T |f^1(s)| ds \right|^2 < \infty$,
- (D4) the generator $f^0 : \Omega_T \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ is such that $(t, \omega) \mapsto f^0(t, \omega, y, z)$ is predictable for all (y, z) , $(y, z) \rightarrow f^0(t, \omega, y, z)$ is continuous for all (t, ω) , and there is an $L_Y \geq 0$ such that, for all (t, ω, y_0, y_1, z) ,

$$|f^0(t, \omega, y_0, z) - f^0(t, \omega, y_1, z)| \leq L_Y |y_0 - y_1|.$$

We let $\Delta\xi := \xi^1 - \xi^0$, and for $s \in [0, T]$,

$$\begin{aligned} \Delta Y_s &:= Y_s^1 - Y_s^0, \\ \Delta Z_s &:= Z_s^1 - Z_s^0, \\ a_s &:= f^1(s) - f^0(s, Y_s^1, Z_s^1), \\ c_s &:= \frac{f^0(s, Y_s^0, Z_s^1) - f^0(s, Y_s^0, Z_s^0)}{|\Delta Z_s|^2} \chi_{\{\Delta Z_s \neq 0\}} \Delta Z_s, \\ \Xi_s &:= |\Delta\xi| + \int_s^T |a_r| dr. \end{aligned}$$

LEMMA 6.34. Assume that $c = (c_t)_{t \in [0, T]} \in \text{BMO}(S_2)$ with $\|c\|_{\text{BMO}(S_2)} \leq \gamma < \infty$, $\lambda_t := \exp(\int_0^t c_s dB_s - \frac{1}{2} \int_0^t |c_s|^2 ds)$ and $p_0 \in (1, \infty)$ such that $\mathcal{RH}_{p'_0}(\lambda) \leq \rho < \infty$ with $1 = (1/p_0) + (1/p'_0)$. Assume $p \in [2, \infty)$ with $p > p_0$ such that

$$\mathbb{E} \left| \int_0^T |\Delta Z_s|^2 ds \right|^{\frac{p}{2}} < \infty.$$

Then there is a $c_{(6.34)} \in (0, \infty)$, depending at most on $(T, L_Y, p, p_0, \gamma, \rho, n)$, such that for all $t \in [0, T]$ one has that

$$\left\| \sup_{s \in [t, T]} |\Delta Y_s| \right\|_p + \left\| \left(\int_t^T |\Delta Z_s|^2 ds \right)^{\frac{1}{2}} \right\|_p \leq c_{(6.34)} \|\Xi_t\|_p.$$

REMARK 6.35. As already mentioned before, Lemma 6.34 continues the work done in [19, Proof of Proposition 2.3], but also the work done in [2, Theorem 5.1]. The main new contribution consist in the fact that using the extension of Fefferman's inequality (Corollary 5.21) we are able to get an L_p - L_p -estimate in contrast to a weaker L_p - L_r -estimate for $r > p$.

PROOF OF LEMMA 6.34. Let $d\mathbb{Q}^* := \lambda_T d\mathbb{Q}$. To distinguish between the integration with respect to \mathbb{Q} and \mathbb{Q}^* , but not to overload the notation, we agree that $\|\cdot\|_p$ always means that we integrate with respect to \mathbb{Q} . By Girsanov's theorem, $(B_s^*)_{s \in [0, T]}$ with $B_s^* := B_s - \int_0^s c_r dr$ is a standard \mathbb{Q}^* -Brownian motion. Now let us fix $t \in [0, T]$ and assume that $\|\Xi_t\|_p < \infty$, otherwise there is nothing to prove. Additionally introducing

$$b_s := \frac{f^0(s, Y_s^1, Z_s^1) - f^0(s, Y_s^0, Z_s^1)}{\Delta Y_s} \chi_{\{\Delta Y_s \neq 0\}},$$

we get that

$$\begin{aligned}
& \Delta Y_t \\
&= \Delta \xi + \int_t^T a_s ds + \int_t^T b_s \Delta Y_s ds + \int_t^T \langle c_s, \Delta Z_s \rangle ds - \int_t^T \Delta Z_s dB_s \\
&= \Delta \xi + \int_t^T a_s ds + \int_t^T b_s \Delta Y_s ds - \int_t^T \Delta Z_s dB_s^*
\end{aligned}$$

where our conditions assure that all terms are well-defined. Because of

$$\mathbb{E}_{\mathbb{Q}^*} \left(\int_0^T |\Delta Z_s|^2 ds \right)^{\frac{1}{2}} \leq \left(\mathbb{E}_{\mathbb{Q}} \lambda_T^{p'} \right)^{\frac{1}{p'}} \left(\mathbb{E}_{\mathbb{Q}} \left(\int_0^T |\Delta Z_s|^2 ds \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} < \infty$$

and the Burkholder-Davis-Gundy inequalities $(\int_0^t \Delta Z_s dB_s^*)_{t \in [0, T]}$ is of class DL and therefore a \mathbb{Q}^* -martingale (see [69, IV.1.7]). Applying Itô's formula implies that

$$e^{\int_0^t b_s ds} \Delta Y_t = e^{\int_0^T b_s ds} \Delta \xi + \int_t^T e^{\int_0^s b_r dr} a_s ds - \int_t^T e^{\int_0^s b_r dr} \Delta Z_s dB_s^*$$

and

$$\Delta Y_t = \mathbb{E}_{\mathbb{Q}^*} \left(e^{\int_t^T b_s ds} \Delta \xi + \int_t^T e^{\int_t^s b_r dr} a_s ds \middle| \mathcal{A}_t \right).$$

Using $p_0 \in (1, p)$ we continue with

$$|\Delta Y_t| \leq e^{(T-t)L_Y} \mathbb{E}_{\mathbb{Q}^*} (\Xi_t | \mathcal{A}_t) \leq e^{(T-t)L_Y} \rho \left(\mathbb{E}_{\mathbb{Q}} (\Xi_t^{p_0} | \mathcal{A}_t) \right)^{\frac{1}{p_0}} \text{ a.s.}$$

By Doob's maximal inequality,

$$(43) \quad \left\| \sup_{s \in [t, T]} |\Delta Y_s| \right\|_p \leq c_{(43)} \|\Xi_t\|_p$$

with $c_{(43)} := e^{(T-t)L_Y} \rho \left(\frac{p}{p-p_0} \right)^{\frac{1}{p_0}}$. Letting

$$\Delta f_s := f^1(s) - f^0(s, Y_s^0, Z_s^0),$$

we also have that

$$|\Delta f_s| \leq |a_s| + |b_s| |\Delta Y_s| + |c_s| |\Delta Z_s| \leq |a_s| + L_Y |\Delta Y_s| + |c_s| |\Delta Z_s|$$

and

$$\begin{aligned}
& \int_t^T |\Delta Y_s \Delta f_s| ds \\
& \leq \int_t^T |\Delta Y_s| [|a_s| + L_Y |\Delta Y_s| + |c_s| |\Delta Z_s|] ds \\
& \leq \sup_{s \in [t, T]} |\Delta Y_s| \int_t^T |a_s| ds + L_Y \int_t^T |\Delta Y_s|^2 ds + \int_t^T [|c_s| |\Delta Y_s| |\Delta Z_s|] ds \\
& \leq \frac{1}{2} \sup_{s \in [t, T]} |\Delta Y_s|^2 + \frac{1}{2} \left[\int_t^T |a_s| ds \right]^2 \\
& \quad + L_Y \int_t^T |\Delta Y_s|^2 ds + \int_t^T [|c_s| |\Delta Y_s| |\Delta Z_s|] ds
\end{aligned}$$

$$\leq \Gamma^2 \sup_{s \in [t, T]} |\Delta Y_s|^2 + \frac{1}{2} \left[\int_t^T |a_s| ds \right]^2 + \int_t^T [|c_s| |\Delta Y_s| |\Delta Z_s|] ds$$

with $\Gamma^2 := \frac{1}{2} + T L_Y$. Now for $S_t(Z)^2 := \int_t^T |\Delta Z_s|^2 ds$ and ${}^*Y_t := \sup_{s \in [t, T]} |\Delta Y_s|$ using Itô's formula, the Burkholder-Davis-Gundy inequalities (8), and Corollary 5.21, we get that

$$\begin{aligned} & \|S_t(Z)\|_p \\ & \leq \left\| \left(|\Delta \xi|^2 + 2 \left| \int_t^T \Delta Y_s \Delta Z_s dB_s \right| + 2 \int_t^T |\Delta Y_s \Delta f_s| ds \right)^{\frac{1}{2}} \right\|_p \\ & \leq \left\| \left(|\Delta \xi|^2 + 2 \left| \int_t^T \Delta Y_s \Delta Z_s dB_s \right| + 2\Gamma^2 {}^*Y_t^2 + \left[\int_t^T |a_s| ds \right]^2 \right. \right. \\ & \quad \left. \left. + 2 \int_t^T [|c_s| |\Delta Y_s| |\Delta Z_s|] ds \right)^{\frac{1}{2}} \right\|_p \\ & \leq \|\Xi_t\|_p + \sqrt{2} \left\| \int_t^T [|c_s| |\Delta Y_s| |\Delta Z_s|] ds \right\|_{\frac{p}{2}}^{\frac{1}{2}} + \sqrt{2} \left\| \int_t^T \Delta Y_s \Delta Z_s dB_s \right\|_{\frac{p}{2}}^{\frac{1}{2}} \\ & \quad + \sqrt{2} \Gamma \|{}^*Y_t\|_p \\ & \leq \|\Xi_t\|_p + \sqrt{2c_{(5.21), \frac{p}{2}}} \|c\|_{\text{BMO}(S_2)}^{\frac{1}{2}} \left\| \left(\int_t^T [|\Delta Y_s| |\Delta Z_s|]^2 ds \right)^{\frac{1}{2}} \right\|_{\frac{p}{2}}^{\frac{1}{2}} \\ & \quad + \sqrt{2\beta_{p/2}} \left\| \left(\int_t^T [|\Delta Y_s| |\Delta Z_s|]^2 ds \right)^{\frac{1}{2}} \right\|_{\frac{p}{2}}^{\frac{1}{2}} + \sqrt{2} \Gamma \|{}^*Y_t\|_p \\ & = \|\Xi_t\|_p + \left[\sqrt{2c_{(5.21), \frac{p}{2}}} \|c\|_{\text{BMO}(S_2)}^{\frac{1}{2}} + \sqrt{2\beta_{p/2}} \right] \times \\ & \quad \times \left\| \left(\int_t^T [|\Delta Y_s| |\Delta Z_s|]^2 ds \right)^{\frac{1}{2}} \right\|_{\frac{p}{2}}^{\frac{1}{2}} + \sqrt{2} \Gamma \|{}^*Y_t\|_p. \end{aligned}$$

Therefore, for $\kappa := \sqrt{2c_{(5.21), \frac{p}{2}}} \gamma + \sqrt{2\beta_{p/2}}$ and $\lambda > 0$ we obtained that

$$\begin{aligned} \|S_t(Z)\|_p & \leq \|\Xi_t\|_p + \kappa \left\| \left(\int_t^T [|\Delta Y_s| |\Delta Z_s|]^2 ds \right)^{\frac{1}{2}} \right\|_{\frac{p}{2}}^{\frac{1}{2}} + \sqrt{2} \Gamma \|{}^*Y_t\|_p \\ & \leq \|\Xi_t\|_p + \kappa \|{}^*Y_t S_t(Z)\|_{\frac{p}{2}}^{\frac{1}{2}} + \sqrt{2} \Gamma \|{}^*Y_t\|_p \\ & \leq \|\Xi_t\|_p + \kappa \left\| \frac{\lambda}{2} {}^*Y_t^2 + \frac{1}{2\lambda} S_t(Z)^2 \right\|_{\frac{p}{2}}^{\frac{1}{2}} + \sqrt{2} \Gamma \|{}^*Y_t\|_p \end{aligned}$$

$$\leq \|\Xi_t\|_p + \kappa \sqrt{\frac{\lambda}{2}} \|*Y_t\|_p + \kappa \sqrt{\frac{1}{2\lambda}} \|S_t(Z)\|_p + \sqrt{2\Gamma} \left\| *Y_t \right\|_p.$$

Choosing $\lambda := 2\kappa^2$ and using (43) gives that

$$\begin{aligned} \|S_t(Z)\|_p &\leq 2\|\Xi_t\|_p + \left[2\kappa^2 + 2\sqrt{2\Gamma} \right] \left\| *Y_t \right\|_p \\ &\leq 2\|\Xi_t\|_p + \left[2\kappa^2 + 2\sqrt{2\Gamma} \right] c_{(43)} \|\Xi_t\|_p \end{aligned}$$

which concludes the proof. \square

6.7. Concluding remarks

We want to conclude this chapter by some connections to the literature not mentioned so far.

6.7.1. Variational estimates. An important problem is the construction of simulation schemes for BSDEs and the study of convergence properties of these schemes. Here quadratic BSDEs are of particular interest, see for example [23]. In certain situations the intuition is confirmed that the rate of convergence of the underlying discretization scheme is upper bounded by the variation of the exact solution to the BSDE (see for example [14, 77, 13]). This is one motivation to find upper bounds for quantities such as $\mathbb{E} \sup_{r \in [s, t]} |Y_r - Y_s|^2$. In relation to simulation schemes, the variation of the Y and Z -processes has been studied in the case of a quadratic generator for example in [46] and [47]. Here (and in most of the articles concerning simulation schemes) the BSDE is assumed to be Markovian and the terminal condition to be a Lipschitz functional of a forward diffusion. For non-Lipschitz terminal conditions the variation of the Y and Z -processes has been also studied in [40] in the Markovian case for Lipschitz BSDEs. This was extended to a class a path-dependent terminal conditions in [35]. In [44, Theorem 2.6] regularity results for non-Markovian Lipschitz BSDEs were proved under assumptions on the Malliavin derivatives up to the second order of ξ and f . In Theorem 2.3 of [44] a condition $M^{2,q}$ is used to investigate the variation of the Y -process of the solution to a BSDE with a random linear generator. The structure of this BSDE yields to an explicit representation of the Y process. The condition $M^{2,q}$ relates to our $\mathbb{B}_p^{\Phi_2}$ spaces via Theorem 4.19. Translated to our setting, the condition $M^{2,q}$ is a condition on the predictable projection of $(D_t \xi)_{t \in [0, T]}$, whereas our condition is a condition on $(D_t \xi)_{t \in [0, T]}$ itself – however the condition in [44] is not a condition on ξ , but on $\xi \rho_T$, where ρ_T is a stochastic exponential.

6.7.2. Existence results. In [24] the existence and uniqueness of solutions to quadratic BSDEs is studied when the terminal condition ξ has a uniformly bounded Malliavin derivative, i.e. $|D \cdot \xi(\cdot)| \leq c$ a.e. which again relates to our spaces $\mathbb{B}_p^{\Phi_2}$. The existence of solutions to some multidimensional quadratic BSDEs, examining as a special case sub-quadratic BSDEs, is considered in [25] under the assumption that the terminal condition is bounded.

6.7.3. Sliceability condition. In [33] the sliceability condition is applied directly to ξ , instead of to $|Z|^\theta$ as in our Theorem 6.4. This is done to consider a new concept of a solution to a BSDE, called *split solution*, to solve quadratic multidimensional BSDEs.

APPENDIX A

Technical Facts

Let $M \neq \emptyset$ be a complete metric space that is *locally σ -compact*, i.e. there exist compact subsets $\emptyset \neq K_1 \subseteq K_2 \subseteq \dots$, such that $\overline{K_n} = K_n$ and $M = \bigcup_{n=1}^{\infty} K_n$. By continuity of a stochastic process $(X_x)_{x \in M} : \Omega \rightarrow \mathbb{R}$ we understand that $x \mapsto X_x(\omega)$ is continuous for all $\omega \in \Omega$.

PROPOSITION A.1. *Let $M \neq \emptyset$ be a complete locally σ -compact metric space and $(X_x)_{x \in M}$ be a continuous process defined on a probability space $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$, and let $(\beta_x)_{x \in M}$ be a stochastic process on a probability space $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$ such that X and β have the same finite-dimensional distributions. Then the following is satisfied:*

- (1) *There exists a continuous process $(Y_x)_{x \in M}$ on $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$, which is a modification of $(\beta_x)_{x \in M}$, i.e. $\mathbb{P}^1(Y_x = \beta_x) = 1$ for all $x \in M$.*
- (2) *If there is another process Y' with this property, then $\mathbb{P}^1(Y_x = Y'_x, x \in M) = 1$.*
- (3) *If $\mathcal{G}^1 \subseteq \mathcal{F}^1$ is a sub- σ -algebra and $D \subseteq M$ dense, such that β_x is \mathcal{G}^1 -measurable for all $x \in D$, then the process Y can be taken to be \mathcal{G}^1 -measurable.*

PROOF. There is a countable set $D_0 = \{a_k : k \geq 1\} \subseteq D$ such that $D_0 \subseteq M$ is dense as well. Taking a sequence $(K_n)_{n=1}^{\infty}$ like in the definition of locally σ -compact we have therefore that $D_0 \cap K_n$ is dense in K_n for all $n = 1, 2, \dots$

(1) and (3): We prove both parts at the same time as (1) is a special case of (3) by taking $D = M$ and $\mathcal{G}^1 = \mathcal{F}^1$. Let K be one of the sets K_n and $A := D_0 \cap K$. Since $x \mapsto X_x$ is continuous on M , it is uniformly continuous on K and A . Hence the set

$$\bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{\substack{d(u,v) < \frac{1}{m} \\ u,v \in A}} \left\{ \omega : |X_u(\omega) - X_v(\omega)| \leq \frac{1}{n} \right\} \in \mathcal{F}^0$$

is of \mathbb{P}^0 -measure one. By the fact that $X \stackrel{d}{=} \beta$, there exists $\Omega_0^1 \in \mathcal{G}^1$ with $\mathbb{P}^1(\Omega_0^1) = 1$ such that $x \mapsto \beta_x(\omega)$ is uniformly continuous on A for all $\omega \in \Omega_0^1$. Since A is dense in K we can define for all $x \in K$ the extension

$$Y_x(\omega) := \begin{cases} \lim_{\substack{x_n \rightarrow x \\ x_n \in A}} \beta_{x_n}(\omega) & : \omega \in \Omega_0^1 \\ 0 & : \omega \in \Omega^1 \setminus \Omega_0^1 \end{cases}.$$

We obtain a \mathcal{G}^1 -measurable continuous process $(Y_x)_{x \in K}$. Take $d \geq 1$, $x_1, \dots, x_d \in K$, and $a_{j,m} \in A$ with $a_{j,m} \rightarrow x_j$ as $m \rightarrow \infty$. Then, for $(t_1, \dots, t_d) \in \mathbb{R}^d$,

$$\begin{aligned} \int_{\Omega^1} e^{i \sum_{j=1}^d t_j Y_{x_j}} d\mathbb{P}^1 &= \lim_{m \rightarrow \infty} \int_{\Omega^1} e^{i \sum_{j=1}^d t_j \beta_{a_{j,m}}} d\mathbb{P}^1 \\ &= \lim_{m \rightarrow \infty} \int_{\Omega^0} e^{i \sum_{j=1}^d t_j X_{a_{j,m}}} d\mathbb{P}^0 \end{aligned}$$

$$= \int_{\Omega^0} e^{i \sum_{j=1}^d t_j X_{x_j}} d\mathbb{P}^0$$

so the finite-dimensional distributions of Y and X coincide. To prove $\mathbb{P}^1(Y_x = \beta_x) = 1$ for all $x \in K$ we check $\mathbb{P}^1(|Y_x - \beta_x| > \epsilon) = 0$ for all $\epsilon > 0$ and all $x \in K$. Let $\epsilon > 0$, $x \in K$, and choose $(x_k)_{k \geq 1} \subseteq A$ such that $x_k \rightarrow_k x$. Then

$$\begin{aligned} \mathbb{P}^1(|Y_x - \beta_x| > \epsilon) &\leq \mathbb{P}^1\left(|Y_x - \beta_{x_k}| > \frac{\epsilon}{2}\right) + \mathbb{P}^1\left(|\beta_x - \beta_{x_k}| > \frac{\epsilon}{2}\right) \\ &= 2\mathbb{P}^0\left(|X_x - X_{x_k}| > \frac{\epsilon}{2}\right) \rightarrow_k 0, \end{aligned}$$

where we used the fact that $Y \stackrel{d}{=} X \stackrel{d}{=} \beta$ and the fact that X is continuous. Thus on any compact $K_n \subseteq M$ we have a continuous \mathcal{G}^1 -measurable process $(Y_x^n)_{x \in K_n}$, that is a modification of $(\beta_x)_{x \in K_n}$. Up to \mathcal{G}^1 -measurable null-sets the construction is consistent in n so that we can construct a \mathcal{G}^1 -measurable continuous process $(Y_x)_{x \in M}$ (where we use $M = \bigcup_{n=1}^\infty K_n$) that is a modification of β .

(2) follows from the separability of M . \square

The following lemma is well-known.

LEMMA A.2. *Let (A, \mathcal{A}) be a measurable space and M be a separable metric space. Assume that $f : M \times A \rightarrow \mathbb{R}$ is such that $f(x, \cdot)$ is \mathcal{A} -measurable for all $x \in M$ and $x \rightarrow f(x, \omega)$ is continuous for all $\omega \in A$. Then f is $\mathcal{B}(M) \otimes \mathcal{A}$ -measurable, where $\mathcal{B}(M)$ is generated by the open sets.*

PROOF. Let $(x_j)_{j \geq 1} \subseteq M$ be a dense set. We define for all $n, j \geq 1$

$$B_j^n := \left\{ x \in M : d(x, x_j) \leq \frac{1}{n} \right\}$$

and obtain a sequence of disjoint sets as follows: $A_1^n := B_1^n$, and $A_k^n := B_k^n \setminus (\bigcup_{j=1}^{k-1} A_j^n)$ for $k = 2, 3, \dots$. Then $M = \bigcup_{k=1}^\infty A_k^n$ for all $n \geq 1$. Now we define $f^n : M \times A \rightarrow \mathbb{R}$ as follows:

$$f^n(x, \omega) := \sum_{j=1}^\infty f(x_j, \omega) 1_{A_j^n}(x).$$

Since $f(x, \cdot)$ is \mathcal{A} -measurable for all $x \in M$ and $A_j^n \in \mathcal{B}(M)$ for all $j, n \geq 1$, it follows that each f^n is $\mathcal{B}(M) \otimes \mathcal{A}$ -measurable. Moreover, for any $(x, \omega) \in M \times A$ we have the pointwise convergence $f^n(x, \omega) \rightarrow f(x, \omega)$ as $n \rightarrow \infty$. This follows from the facts

$$|f^n(x, \omega) - f(x, \omega)| = |f(x_{j(n,x)}, \omega) - f(x, \omega)|,$$

and $d(x_{j(n,x)}, x) \leq \frac{1}{n} \rightarrow_n 0$, where $j(n, x)$ is the index such that $x \in A_{j(n,x)}^n$. \square

LEMMA A.3. *Let $f \in L_1([0, T])$ be non-negative. Then*

$$\sup_{0 \leq a < b \leq T} \frac{1}{b-a} \int_a^b f(t) dt = \text{esssup}_{t \in [0, T]} f.$$

PROOF. The inequality

$$\sup_{0 \leq a < b \leq T} \frac{1}{b-a} \int_a^b f(t) dt \leq \text{esssup}_{t \in [0, T]} f.$$

is obvious. According to [72, Theorem 3.3.8] there exists a Borel set $A \subseteq [0, T]$ with $\lambda(A) = T$ and $0 \leq a_n^s \leq s \leq b_n^s \leq T$ with $0 < b_n^s - a_n^s \rightarrow_n 0$ for $s \in A$, such that

$$\lim_n \frac{1}{b_n^s - a_n^s} \int_{a_n^s}^{b_n^s} f(t) dt = f(s)$$

for all $s \in A$. Hence,

$$f(s) \leq \sup_{0 \leq a < b \leq T} \frac{1}{b - a} \int_a^b f(t) dt$$

for all $s \in A$. □

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, H be a separable Hilbert space with $H \neq \{0\}$, and $(g_h)_{h \in H}$ be an iso-normal family of Gaussian random variables $g_h : \Omega \rightarrow \mathbb{R}$. Assume that

$$\mathcal{F} = \sigma(g_h : h \in H) \vee \mathcal{N}$$

where \mathcal{N} are the null-sets from \mathcal{F} . Let $(e_k)_{k \in I}$ be an orthonormal basis of H with $I = \{1, \dots, d\}$ or $I = \{1, 2, \dots\}$. Then

$$\mathcal{F} = \sigma(g_{e_k} : k \in I) \vee \mathcal{N}.$$

We recall that $D : \mathbb{D}_{1,2} \rightarrow L_2^H$ is a closed operator (see [59, Proposition 1.2.1]). Assume that $\varphi_n : \mathbb{R} \rightarrow [0, \infty) \in C_0^\infty$ such that $\varphi_n(x) = 0$ for $x \leq 0$ and $x \geq 1/n$ and that $\int_{\mathbb{R}} \varphi_n(x) dx = 1$. Defining $\psi_n(y) := \int_{-\infty}^y \varphi_n(x) dx$, we get $\psi_n(x) = 0$ for $x \leq 0$, $\psi_n(x) = 1$ if $x \geq 1/n$, and $0 \leq \psi_n(x) \leq 1$. Finally, set

$$L_n(y) := \int_{-\infty}^y \psi_n(x) dx$$

so that $L_n'(x) = \psi_n(x) \rightarrow_n \chi_{(0, \infty)}(x)$ and

$$0 \leq x - L_n(x) \leq \frac{1}{n}$$

for $x \geq 0$ whereas $L_n(x) = 0$ for $x \leq 0$. Given $\xi \in \mathbb{D}_{1,2}$ we get that $|\xi^+ - L_n(\xi)| \leq 1/n$ and $L_n'(\xi) D\xi \rightarrow \chi_{(0, \infty)}(\xi) D\xi$ in L_2^H . Hence $\xi^+ \in \mathbb{D}_{1,2}$ with

$$D\xi^+ = \chi_{(0, \infty)}(\xi) D\xi$$

and, for $L > 0$,

$$\begin{aligned} D(\xi \vee (-L)) &= D((\xi + L)^+ - L) \\ &= \chi_{(0, \infty)}(\xi + L) D(\xi + L) = \chi_{(-L, \infty)}(\xi) D(\xi). \end{aligned}$$

From this we get

$$\begin{aligned} D(\xi \wedge L) &= -D((-\xi) \vee (-L)) \\ &= -\chi_{(-L, \infty)}(-\xi) D(-\xi) = \chi_{(-\infty, L)}(\xi) D(\xi). \end{aligned}$$

Finally,

$$\begin{aligned} D((\xi \vee (-L)) \wedge L) &= \chi_{(-\infty, L)}(\xi \vee (-L)) D((\xi \vee (-L)) \wedge L) \\ &= \chi_{(-\infty, L)}(\xi \vee (-L)) \chi_{(-L, \infty)}(\xi) D(\xi) \\ &= \chi_{(-L, L)}(\xi) D(\xi). \end{aligned}$$

PROPOSITION A.4. *Let H be a separable Hilbert space with an orthonormal basis $(e_k)_{k \in I}$, where $I = \{1, \dots, d\}$ or $I = \{1, 2, \dots\}$, let $(g_h)_{h \in H}$, $g_h : \Omega \rightarrow \mathbb{R}$, be an iso-normal family of Gaussian random variables defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathcal{F} = \sigma(g_h : h \in H) \vee \mathcal{N}$ with \mathcal{N} being the null-sets of $(\Omega, \mathcal{F}, \mathbb{P})$. Let $p \in [2, \infty)$, $\varepsilon > 0$, and $\xi \in \mathbb{D}_{1,2} \cap L_p$ such that $D\xi \in L_p^H$. Then there exist $n \geq 1$ and a bounded $f_n \in C^\infty(\mathbb{R}^n)$ such that all derivatives are bounded (where the bound may depend on the order of the derivative) such that for $\xi_0 := f_n(g_{e_1}, \dots, g_{e_n})$ one has*

$$\|\xi - \xi_0\|_p^p + \|D\xi - D\xi_0\|_{L_p^H}^p < \varepsilon^p.$$

PROOF. (a) Reduction to $\dim(H) < \infty$ in the case $\dim(H) = \infty$: Let $\mathcal{H}_n := \sigma(g_{e_1}, \dots, g_{e_n})$. By martingale convergence it follows that

$$\lim_n \xi_n := \lim_n \mathbb{E}(\xi | \mathcal{H}_n) = \xi \text{ a.s. and in } L_p.$$

For $n \in I$ let $P_n : H \rightarrow \text{span}\{e_1, \dots, e_n\} \subseteq H$ be the orthogonal projection. Then

$$\begin{aligned} \|D\xi - D\xi_n\|_{L_p^H} &= \|P_n D\xi - D\xi_n + (I - P_n)D\xi\|_{L_p^H} \\ &\leq \|P_n D\xi - D\xi_n\|_{L_p^H} + \|(I - P_n)D\xi\|_{L_p^H}. \end{aligned}$$

By dominated convergence,

$$\lim_n \|(I - P_n)D\xi\|_{L_p^H} = 0.$$

On the other hand, using $D\xi_n = P_n \mathbb{E}(D\xi | \mathcal{H}_n)$ we get

$$\|P_n D\xi - D\xi_n\|_{L_p^H} = \|P_n D\xi - P_n \mathbb{E}(D\xi | \mathcal{H}_n)\|_{L_p^H} \leq \|D\xi - \mathbb{E}(D\xi | \mathcal{H}_n)\|_{L_p^H}$$

that converges to zero as $n \rightarrow \infty$ because $\mathcal{F} = \bigvee_{n \geq 1} \mathcal{H}_n \vee \mathcal{N}$ and because of known facts about Banach space valued closable martingales. Summing up, we obtain

$$\lim_n \left[\|\xi - \xi_n\|_p^p + \|D\xi - D\xi_n\|_{L_p^H}^p \right] = 0.$$

(b) Reduction to a bounded ξ : For $L \geq 1$ define the truncation function $\psi_L : \mathbb{R} \rightarrow \mathbb{R}$ by $\psi_L(x) := (x \vee (-L)) \wedge L$. Then

$$\lim_{L \rightarrow \infty} \|\xi_n - \psi_L(\xi_n)\|_p = 0$$

where ξ_n is an approximation obtained by (a) or we take $\xi_n = \xi$ in case $\dim(H) < \infty$. Moreover, $\chi_{(-L, L)}(\xi_n) D\xi_n$ is a representative of $D(\psi_L(\xi_n))$, so that

$$\lim_{L \rightarrow \infty} \|D\xi_n - D(\psi_L(\xi_n))\|_{L_p^H} = 0$$

as well. Consequently, for all $\varepsilon > 0$ there are $n, L \geq 1$ such that

$$\|\xi - \psi_L(\xi_n)\|_p^p + \|D\xi - D(\psi_L(\xi_n))\|_{L_p^H}^p < \varepsilon^p.$$

(c) Reduction to the smooth case: By the factorization theorem we can write

$$\psi_L(\xi_n) = f_n(g_{e_1}, \dots, g_{e_n}) \in \mathbb{D}_{1,2}$$

for a bounded Borel function $f_n : \mathbb{R}^n \rightarrow \mathbb{R}$ where we suppress L in the following. Let $F_n : [0, 1) \times \mathbb{R}^n \rightarrow \mathbb{R}$ be the solution of the backward heat equation with terminal condition f_n so that

$$\lim_{t \rightarrow 1} F_n(t, B_t^n) = f_n(B_1^n) \quad \text{and} \quad \lim_{t \rightarrow 1} \nabla F_n(t, B_t^n) = Df_n(B_1^n)$$

in L_p and $L_p^{\mathbb{R}^n}$, respectively, and a.s., where $(B_t^n)_{t \in [0,1]}$ is an n -dimensional standard Brownian motion. But this implies also that

$$\lim_{t \rightarrow 1} F_n(t, \sqrt{t}B_1^n) = f_n(B_1^n) \quad \text{and} \quad \lim_{t \rightarrow 1} \nabla F_n(t, \sqrt{t}B_1^n) = Df_n(B_1^n)$$

in L_p and $L_p^{\mathbb{R}^n}$, respectively. This can be seen from the estimate

$$\begin{aligned} \|F_n(t, \sqrt{t}B_1^n) - f_n(B_1^n)\|_p^p &= \mathbb{E}|\tilde{\mathbb{E}}f_n(\sqrt{t}B_1^n + \tilde{B}_{1-t}^n) - f_n(B_1^n)|^p \\ &\leq \mathbb{E}\tilde{\mathbb{E}}|f_n(\sqrt{t}B_1^n + \tilde{B}_{1-t}^n) - f_n(B_1^n)|^p \\ &= \mathbb{E}\tilde{\mathbb{E}}|f_n(B_{\sqrt{t}}^n + \tilde{B}_{1-\sqrt{t}}^n) - f_n(B_1^n)|^p \end{aligned}$$

so that

$$\|F_n(t, \sqrt{t}B_1^n) - f_n(B_1^n)\|_p \leq 2\|F_n(\sqrt{t}, B_{\sqrt{t}}^n) - f_n(B_1^n)\|_p \rightarrow 0$$

as $t \rightarrow 1$. The fact we used here is that $(F_n(t, B_t^n))_{t \in [0,1]}$ is a martingale. As $(\nabla F_n(t, B_t^n))_{t \in [0,1]}$ is a martingale as well, where we agree about $Df_n =: \nabla F_n(1, \cdot)$, the same computation yields to

$$\|\nabla F_n(t, \sqrt{t}B_1^n) - Df_n(B_1^n)\|_{L_p^H} \leq 2\|\nabla F_n(\sqrt{t}, B_{\sqrt{t}}^n) - Df_n(B_1^n)\|_{L_p^H} \rightarrow 0$$

as $t \rightarrow 1$. Letting $f_{n,t} := F_n(t, \sqrt{t} \cdot)$ for $t \in [0, 1)$, we get that

$$Df_{n,t}(g_{e_1}, \dots, g_{e_n}) = \sqrt{t} \sum_{k=1}^n \frac{\partial}{\partial x_k} F_n(t, \sqrt{t}(g_{e_1}, \dots, g_{e_n})) e_k$$

because $f_{n,t} \in C_1^b(\mathbb{R}^n) \cap C^b(\mathbb{R}^n)$, and therefore

$$\begin{aligned} &\|Df_n(g_{e_1}, \dots, g_{e_n}) - Df_{n,t}(g_{e_1}, \dots, g_{e_n})\|_{L_p^H} \\ &\leq \left\| Df_n(g_{e_1}, \dots, g_{e_n}) - \sum_{k=1}^n \frac{\partial}{\partial x_k} F_n(t, \sqrt{t}(g_{e_1}, \dots, g_{e_n})) e_k \right\|_{L_p^H} \\ &\quad + (1 - \sqrt{t}) \left\| \sum_{k=1}^n \frac{\partial}{\partial x_k} F_n(t, \sqrt{t}(g_{e_1}, \dots, g_{e_n})) e_k \right\|_{L_p^H} \\ &\leq \left\| Df_n(g_{e_1}, \dots, g_{e_n}) - \sum_{k=1}^n \frac{\partial}{\partial x_k} F_n(t, \sqrt{t}(g_{e_1}, \dots, g_{e_n})) e_k \right\|_{L_p^H} \\ &\quad + (1 - \sqrt{t}) \|Df_n(g_{e_1}, \dots, g_{e_n})\|_{L_p^H}. \end{aligned}$$

Summarizing,

$$\begin{aligned} \lim_{t \rightarrow 1} [\|f_{n,t}(g_{e_1}, \dots, g_{e_n}) - f_n(g_{e_1}, \dots, g_{e_n})\|_p^p \\ + \|Df_n(g_{e_1}, \dots, g_{e_n}) - Df_{n,t}(g_{e_1}, \dots, g_{e_n})\|_{L_p^H}^p] = 0. \quad \square \end{aligned}$$

LEMMA A.5 (Stein's martingale inequality, [55] and cf. [67, Theorem 3.2]). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $p \in (1, \infty)$ and let $(\mathcal{G}_k)_{k=1}^n$ be an increasing sequence of sub- σ -algebras of \mathcal{F} . Then one has*

$$\left\| \left(\sum_{k=1}^n |\mathbb{E}(f_k | \mathcal{G}_k)|^2 \right)^{\frac{1}{2}} \right\|_p \leq c_p \left\| \left(\sum_{k=1}^n |f_k|^2 \right)^{\frac{1}{2}} \right\|_p$$

for all $f_1, \dots, f_n \in L_p$ where the constant $c_p > 0$ depends at most on p .

Note that by grouping the random variables in an appropriate way in Stein's inequality, we can also assume that f_1, \dots, f_n are random vectors with values in ℓ_2^N , whereas the constant $c_p > 0$ does not enlarge.

LEMMA A.6. *For $p \in (1, \infty)$ assume a stochastic process $a = (a_t)_{t \in [0,1]}$ with values in ℓ_2^N that has left-continuous paths and satisfies $\mathbb{E} \sup_t |a_t|^p < \infty$. Suppose a filtration $(\mathcal{H}_t)_{t \in [0,1]}$ and an $(\mathcal{H}_t)_{t \in [0,1]}$ -adapted process $(b_t)_{t \in [0,1]}$ with values in ℓ_2^N and $\mathbb{E}|b_t|^p < \infty$ for all $t \in [0,1]$ that has left-continuous paths and such that $b_t = \mathbb{E}(a_t | \mathcal{H}_t)$ a.s. for $t = k/2^n$ with $n = 0, 1, 2, \dots$ and $k = 0, \dots, 2^n - 1$. Then one has that*

$$\left\| \left(\int_0^1 |b_t|^2 dt \right)^{\frac{1}{2}} \right\|_p \leq c_{(A.5)} \left\| \left(\int_0^1 |a_t|^2 dt \right)^{\frac{1}{2}} \right\|_p$$

where $c_{(A.5)} > 0$ is taken from Lemma A.5.

PROOF. Let $t_k^n := \frac{k}{2^n}$ for $n \geq 0$ and $k = 0, \dots, 2^n - 1$. Then it follows from Lemma A.5 that

$$\left\| \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n) |\mathbb{E}(a_{t_k^n} | \mathcal{H}_{t_k^n})|^2 \right)^{\frac{1}{2}} \right\|_p \leq c_{(A.5)} \left\| \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n) |a_{t_k^n}|^2 \right)^{\frac{1}{2}} \right\|_p.$$

Applying twice Fatou's lemma on the left-hand side, we derive

$$\left\| \left(\int_0^1 |b_t|^2 dt \right)^{\frac{1}{2}} \right\|_p \leq c_{(A.5)} \liminf_n \left\| \left(\sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n) |a_{t_k^n}|^2 \right)^{\frac{1}{2}} \right\|_p$$

and we can conclude by dominated convergence. \square

LEMMA A.7. *Let $p \in (1, \infty)$, $N \geq 1$ and $f : \mathbb{R}^N \rightarrow \mathbb{R} \in C^\infty$ where $\|D^\alpha f\|_\infty < \infty$ for all multi-indices α . Let γ_N be the standard Gaussian measure on \mathbb{R}^N . Then one has*

$$\left\| f - \int_{\mathbb{R}^N} f d\gamma_N \right\|_{L_p(\gamma_N)} \leq c_{(A.7)} \|\nabla f\|_{L_p(\gamma_N)}$$

where the constant $c_{(A.7)} > 0$ depends on p only.

PROOF. Let $B = (B_t)_{t \in [0,1]}$ be an N -dimensional standard Brownian motion on a complete probability space (M, Σ, μ) with the augmented natural filtration $(\mathcal{G}_t)_{t \in [0,1]}$ and that $\Sigma = \mathcal{G}_1$. Let

$$F(t, x) := \mathbb{E}f(x + B_{1-t})$$

so that, by Itô's formula,

$$f(B_1) - \mathbb{E}f(B_1) = \int_0^1 \nabla F(t, B_t) dB_t,$$

and, by the Burkholder-Davis-Gundy inequalities,

$$\|f(B_1) - \mathbb{E}f(B_1)\|_p \leq c_p \left\| \left(\int_0^1 |\nabla F(t, B_t)|^2 dt \right)^{\frac{1}{2}} \right\|_p$$

and we can conclude with Lemma A.6 by $a_t \equiv \nabla f(B_1)$, $b_t := \nabla F(t, B_t)$ and $\mathcal{H}_t = \mathcal{G}_t$. \square

We call a function $h : \Omega \rightarrow \mathbb{R}$ a Π -step-function, where $\Pi \subseteq 2^\Omega$ is non-empty system of subsets, provided that $h = \sum_{k=1}^n \alpha_k \chi_{A_k}$ for some $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ and $A_1, \dots, A_n \in \Pi$.

THEOREM A.8. *Let Ω be a non-empty set and Π be a system of subsets of Ω such that*

- (i) $A, B \in \Pi$ implies $A \cap B \in \Pi$,
- (ii) $\Omega \in \Pi$.

Let $p \in [1, \infty)$ and $\mathcal{F} := \sigma(\Pi)$. Then for all $f \in L_p(\Omega, \mathcal{F}, \mathbb{P})$ there are Π -step-functions $f_n : \Omega \rightarrow \mathbb{R}$ such that $\lim_n \|f - f_n\|_p = 0$.

PROOF. Let $\mathcal{M} := \{\chi_A : A \in \Pi\}$ so that $\mathcal{F} = \sigma(\Pi) = \sigma(\mathcal{M})$. Let \mathcal{H} be the set of all bounded measurable $f : \Omega \rightarrow \mathbb{R}$ such that there exist Π -step-functions $h_k : \Omega \rightarrow \mathbb{R}$ with $\lim_k \|f - h_k\|_p = 0$. Then \mathcal{H} and \mathcal{M} satisfy the assumptions of the monotone class theorem (see [65, p. 7]). Hence any bounded \mathcal{F} -measurable function can be approximated in L_p by Π -step-functions. Our assertion follows by one more approximation obtained by truncation of a general element of L_p . \square

THEOREM A.9. *Let $X = (X_t)_{t \in [0, T]}$, $T > 0$, $X_t : \Omega \rightarrow \mathbb{R}^d$, be a stochastic process such that all families $(X_{t_i^k}^k - X_{t_{i-1}^k}^k)_{k=1, i=1}^{d, N_k}$ with*

$$0 = t_0^k < \dots < t_{N_k}^k = T \quad \text{and} \quad N_k \geq 1$$

are independent, $\mathcal{F} := \sigma(X)$, and $p \in [1, \infty)$. Then the following holds:

- (i) *The linear span of*

$$\prod_{k=1}^d \prod_{i=1}^{N_k} \chi_{\{X_{t_i^k}^k - X_{t_{i-1}^k}^k \in (a_i^k, b_i^k)\}},$$

where for $N_k = 0$ the corresponding product is replaced by 1 and for $N_k \geq 1$ we have $-\infty < a_i^k < b_i^k < \infty$ and $0 \leq t_{i-1}^k < t_i^k \leq T$, is dense in $L_p(\Omega, \mathcal{F}, \mathbb{P})$.

- (ii) *If X is the d -dimensional standard Brownian motion, then the linear span of*

$$\prod_{k=1}^d \prod_{i=1}^{N_k} (X_{t_i^k}^k - X_{t_{i-1}^k}^k)$$

is dense in $L_2(\Omega, \mathcal{F}, \mathbb{P})$, where for $N_k = 0$ the corresponding product is replaced by 1 and for $N_k \geq 1$ the intervals $(t_{i-1}^k, t_i^k]$, $i = 1, \dots, N_k$, are pair-wise disjoint for any fixed k .

PROOF. (i) The system Π consisting of Ω and all possible finite intersections of $\{X_t^k - X_s^k \in (a, b)\}$ with $k \in \{1, \dots, d\}$, $0 \leq s < t \leq T$, and $-\infty < a < b < \infty$, satisfies (i) and (ii) of Theorem A.8 and $\mathcal{F} = \sigma(\Pi)$. Therefore assertion (i) follows from the same Theorem A.8.

- (ii) By step (i) the random variables of form

$$\xi = f\left(\frac{X_{t_1} - X_{t_0}}{\sqrt{t_1 - t_0}}, \dots, \frac{X_{t_n} - X_{t_{n-1}}}{\sqrt{t_n - t_{n-1}}}\right),$$

where $n \geq 1$, $0 = t_0 < \dots < t_n = T$ and $f : \mathbb{R}^{nd} \rightarrow \mathbb{R}$ is a bounded Borel function, are dense in $L_2(\Omega, \mathcal{F}, \mathbb{P})$. Exploiting the orthonormal basis of Hermite functions of $L_2(\mathbb{R}^{nd}, \gamma_{nd})$ we can approximate ξ by polynomials in $(X_{t_i}^k - X_{t_{i-1}}^k)$

where $k = 1, \dots, d$ and $i = 1, \dots, n$. It remains to approximate $(X_b^k - X_a^k)^l$ for $l \geq 2$, $k \in \{1, \dots, d\}$ and $0 \leq a < b \leq T$ by

$$\sum_{\substack{i_1, \dots, i_l \in \{1, \dots, N\} \\ \text{distinct}}} \left(X_{a+i_1 \frac{b-a}{N}}^k - X_{a+(i_1-1) \frac{b-a}{N}}^k \right) \cdots \left(X_{a+i_l \frac{b-a}{N}}^k - X_{a+(i_l-1) \frac{b-a}{N}}^k \right)$$

and $N \rightarrow \infty$. \square

The following lemma can be proved by the generalized Clark-Ocone formula from [60, Proposition A.1]. For completeness we include an argument based on a periodic time-shift of the Brownian motion.

LEMMA A.10. *Let $p \in [2, \infty)$, $\xi = \sum_{k=0}^{\infty} I_k(f_k) \in \mathbb{D}_{1,2} \cap L_p(\Omega, \mathcal{F}, \mathbb{P})$ with symmetric kernels f_k , and $b \in (0, T]$. Then there are measurable processes $(\mu_t^b(i))_{t \in [0, b]}$, $i = 1, \dots, d$, such that for all $a \in [0, b)$ one has*

$$(1) \quad \|\xi - \mathbb{E}(\xi | \mathcal{G}_a^b)\|_p \sim_{\kappa_p} \left\| \left(\int_a^b |\mu_r^b|^2 dr \right)^{\frac{1}{2}} \right\|_p, \text{ where } \kappa_p \geq 1 \text{ depends on } p \text{ only,}$$

$$(2) \text{ and that}$$

$$\begin{aligned} & \int_{(a,b]} \mathbb{E} |\mu_r^b(i) - D(r, i) \xi|^2 dr \\ &= \int_{(a,b]} \sum_{k=1}^{\infty} k^2 (k-1)! \|f_k((r, i), \cdot) [\chi_{((0,r] \cup (b,T))^{k-1}} - 1]\|_{L_2}^2 dr. \end{aligned}$$

PROOF. We represent our Wiener space by a different Brownian motion, obtained by a permutation of the original one. For this purpose we let

$$W_t^b := \begin{cases} W_{b+t} - W_b & : t \in [0, T-b] \\ W_{t-T+b} + W_T - W_b & : t \in [T-b, T] \end{cases}$$

and obtain a standard Brownian motion (as Gaussian process). We have that $\sigma(W_t^b : t \in [0, T]) = \sigma(W_t : t \in [0, T])$ and $\mathcal{G}_t^b = \mathcal{F}_{T-b+t}^{W^b}$. The symmetric kernels f_n for the chaos decompositions with respect to W may be transformed to W^b as

$$(44) \quad f_n^b((t_1, i_1), \dots, (t_n, i_n)) = f_n(((\varphi^b)^{-1}(t_1), i_1), \dots, ((\varphi^b)^{-1}(t_n), i_n))$$

where $\varphi^b(t) := t + (T-b)$ for $t \in (0, b]$ and $\varphi^b(t) := t - b$ for $t \in (b, T]$. Now we get that

$$\xi - \mathbb{E}(\xi | \mathcal{G}_a^b) = \xi - \mathbb{E}(\xi | \mathcal{F}_{T-b+a}^{W^b}).$$

Let $\xi = \sum_{n=0}^{\infty} I_n(f_n^b)$ the chaos decomposition with respect to W^b where the kernels are obtained from the representation in terms of W by formula (44). Exploiting the representation property on the Wiener space, we find progressively measurable (with respect to the augmentation of the natural filtration $(\mathcal{F}_t^{W^b})_{t \in [0, T]}$ of $(W_t^b)_{t \in [0, T]}$) processes $(\lambda_t^b(i))_{t \in [0, T]}$, $i = 1, \dots, d$, satisfying $\mathbb{E} \int_0^T |\lambda_t^b|^2 dt < \infty$ and

$$\xi = \mathbb{E}\xi + \int_{(0, T]} \lambda_t^b dW_t^b \text{ a.s.}$$

Then the processes $(\mu_t^b(i))_{t \in [0, b]}$ are defined by

$$\mu_r^b(i) := \lambda_{T-b+r}^b(i).$$

By the Burkholder-Davis-Gundy inequalities we get that

$$\begin{aligned}
\|\xi - \mathbb{E}(\xi|\mathcal{G}_a^b)\|_p &= \|\xi - \mathbb{E}(\xi|\mathcal{F}_{T-b+a}^{W^b})\|_p \\
&\sim \kappa_p \left\| \left(\int_{T-b+a}^T |\lambda_r^b|^2 dr \right)^{\frac{1}{2}} \right\|_p = \left\| \left(\int_a^b |\mu_r^b|^2 dr \right)^{\frac{1}{2}} \right\|_p.
\end{aligned}$$

This proves part (1). Regarding part (2) it is sufficient to prove the equality for ξ from a dense subset of $\mathbb{D}_{1,2}$. So we may assume $\xi = \sum_{k=1}^N I_k(f_k)$, $N \geq 1$, with symmetric f_k that are constant on dyadic cuboids of side-length $T/2^L$, $L \geq 1$, and vanish on diagonal cuboids (where at least two edges coincide). For those ξ we have the explicit formula

$$\lambda_t^b(i) = \sum_{k=1}^N k I_{k-1}^b(f_k^b((t, i), \cdot) \chi_{(0, t]^{k-1}})$$

where we chose the canonical representatives on the right-hand side. In this case one can directly check part (2). \square

Acknowledgment: We would like to thank Christian Bender, Christel Geiss, David Nualart, Adrien Richou, and Alexander Steinicke for helpful discussions.

Bibliography

- [1] H. Airault, J. Ren and X. Zhang: Smoothness of local times of semimartingales. *C.R. Acad. Sci. Paris* 330:719-724, 2000.
- [2] S. Ankirchner, P. Imkeller and G. Dos Reis: Classical and variational differentiability of BSDEs with quadratic growth. *Electronic Journal Prob.* 12:1418-1453, 2007.
- [3] R. Avikainen: On irregular functionals of SDEs and the Euler scheme. *Finance and Stochastics* 13:381-401, 2009.
- [4] R. Bañuelos and A.G. Bennett: Paraproducts and Commutators of Martingale Transforms. *Proc. AMS* 103:1226-1234, 1988.
- [5] P. Barrieu and N. El Karoui: Monotone stability of quadratic semimartingales with applications to unbounded general quadratic BSDEs. *Ann. Prob.* 41:1831-1863, 2013.
- [6] H. Bauer: Measure and Integration Theory. de Gruyter, 2001.
- [7] F. Baumgartner and S. Geiss. Permutation Invariant Functionals of Lévy Processes. ArXiv 1407.3645, in revision for *Trans. AMS*.
- [8] C. Bennett and R. Sharpley: Interpolation of Operators Academic Press, 1988.
- [9] J. Bergh and J. Löfström: Interpolation spaces. An Introduction. Springer, 1976.
- [10] A. Bhattacharyya: On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.* 35:99-109, 1943.
- [11] J.M. Bismut: Théorie probabiliste du contrôle des diffusions. *Mem. AMS* 176, 1973.
- [12] A. Bonami and D. Lépine: Fonction maximale et variation quadratique des martingales en présence d'un poids. Séminaire de Probabilités XIII, Univ. de Strasbourg, Lecture Notes in Mathematics 721, pp. 294-306, Springer, 1979.
- [13] B. Bouchard, R. Elie and N. Touzi: Discrete-time approximation of BSDEs and probabilistic schemes for fully nonlinear PDEs. *Radon Series Comp. Appl. Math* 8, 1-34, 2009
- [14] B. Bouchard and N. Touzi: Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stoch. Proc. Appl.* 111:175-206, 2004.
- [15] B. Boufoussi: Régularité du temps local Brownien dans les espaces de Besov-Orlicz. *Studia Math.* 118(2):145-156, 1996.
- [16] B. Boufoussi and B. Roynette: Le temps local Brownien appartient p.s. à l'espace de Besov $\mathbb{B}_{p,\alpha}^{1/2}$. *C.R. Acad. Sci. Paris Série I* 316:843-848, 1993.
- [17] P. Briand and F. Confortola: BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces. *Stoch. Proc. Appl.* 118:818-838, 2008.
- [18] P. Briand, B. Delyon, Y. Hu, E. Pardoux and L. Stoica: L_p solutions of backward stochastic differential equations. *Stoch. Proc. Appl.* 108:109-129, 2003.
- [19] P. Briand and R. Elie: A simple constructive approach to quadratic BSDEs with or without delay. *Stoch. Proc. Appl.* 123:2921-2939, 2013.
- [20] P. Briand and Y. Hu: BSDEs with quadratic growth and unbounded terminal value. *Probab. Theory Relat. Fields* 136:604-618, 2006.
- [21] P. Briand and Y. Hu: Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Relat. Fields* 141:543-567, 2008.
- [22] P. Briand, J.-P. Lepeltier and J.S. Martin: One-dimensional backward stochastic differential equations whose coefficient is monotone in y and non-Lipschitz in z . *Bernoulli* 13:80-91, 2007.
- [23] J.-F. Chassagneux and A. Richou: Numerical simulation of quadratic BSDEs. *Ann. Appl. Prob.* 26:262-304, 2016.
- [24] P. Cheridito and K. Nam: BSDEs with terminal conditions that have bounded Malliavin derivative. *J. Funct. Anal.* 266:1257-1285, 2014.

- [25] P. Cheridito and K. Nam: Multidimensional quadratic and subquadratic BSDEs with special structure. *Stochastics: An International Journal of Probability and Stochastic Processes* 87:871-884, 2015.
- [26] F. Coquet, Y. Hu, J. Memin and S. Peng: Filtration-consistent nonlinear expectations and related g -expectations. *Prob. Theory Relat. Fields* 123:1-27, 2002.
- [27] F. Delbaen, Y. Hu and A. Richou: On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions. *Ann. Inst. Henri Poincaré Probab. Stat.* 47:559-574, 2011.
- [28] F. Delbaen, Y. Hu and A. Richou: On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions: the critical case. *Discrete Contin. Dyn. Syst.* 35:5273-5283, 2015.
- [29] F. Delbaen and S. Tang: Harmonic analysis of stochastic equations and backward stochastic differential equations. *Prob. Theory Relat. Fields* 146:291-336, 2010.
- [30] N. El Karoui, S. Peng and M.C. Quenez: Backward stochastic differential equations in finance. *Mathematical Finance* 7:1-71, 1997.
- [31] M. Emery: Stabilité des solutions des équations différentielles stochastiques; applications aux intégrales multiplicatives stochastiques. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* 41:241-262, 1978.
- [32] M. Emery: Équations différentielles stochastiques lipschitziennes: étude de la stabilité. Séminaire de Probabilités XIII. Lecture Notes in Mathematics 721, pp. 281-293. Springer, 1979.
- [33] C. Frei: Splitting multidimensional BSDEs and finding local equilibria. *Stoch. Proc. Appl.* 124:2654-2671, 2014.
- [34] A.M. Garsia: Martingale Inequalities. W.A. Benjamin, 1973.
- [35] C. Geiss, S. Geiss and E. Gobet: Generalized fractional smoothness and L_p -variation of BSDEs with non-Lipschitz terminal condition. *Stoch. Proc. Appl.* 122:2078-2116, 2012.
- [36] C. Geiss and A. Steinicke. L_2 -variation of Lévy driven BSDEs with non-smooth terminal conditions. *Bernoulli*, 22(2):995-1025, 2016.
- [37] S. Geiss: Weighted BMO and discrete time hedging within the Black-Scholes model. *Prob. Theory Related Fields* 132:39-73, 2005.
- [38] S. Geiss and M. Hujo: Interpolation and approximation in $L_2(\gamma)$. *J. Approx. Theory* 144:213-232, 2007.
- [39] S. Geiss and A. Toivola: On fractional smoothness and L_p -approximation on the Gaussian space. *Ann. Prob.* 43:605-638, 2015.
- [40] E. Gobet and A. Makhlof: L_2 -time regularity of BSDEs with irregular terminal functions. *Stoch. Proc. Appl.* 120:1105-1132, 2010.
- [41] P. Grandits: On a conjecture of Kazamaki. Séminaire de Probabilités XXX. Lecture Notes in Mathematics 1626, pp. 357-360. Springer, 1996.
- [42] F. Hirsch: Lipschitz functions and fractional Sobolev spaces. *Potential Analysis* 11:415-429, 1999.
- [43] Y. Hu, P. Imkeller and M. Müller: Utility maximization in incomplete markets. *Ann. Appl. Prob.* 15:1691-1712, 2005.
- [44] Y. Hu, D. Nualart and X. Song: Malliavin Calculus for backward stochastic differential equations and applications to numerical solutions. *Ann. Appl. Prob.* 21:2379-2423, 2011.
- [45] P. Imkeller: Malliavin's calculus and applications in stochastic control and finance. *IM PAN Lecture Notes*, Warsaw 2009.
- [46] P. Imkeller and G. dos Reis: Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *Stoch. Proc. Appl.* 120:348-379, 2010.
- [47] P. Imkeller, G. dos Reis, and J. Zhang: Results on numerics for FBSDE with drivers of quadratic growth. *Contemporary Quantitative Finance*, Springer Berlin, 2010.
- [48] M. Izumisawa and N. Kazamaki: Weighted norm inequalities for martingales. *Tôhoku Math. Journal* 29:115-124, 1977.
- [49] I. Karatzas and S. Shreve: Brownian Motion and Stochastic Calculus. 2nd Edition. Springer, 1991.
- [50] N. Kazamaki: Continuous Exponential Martingales and BMO. Lecture Notes in Mathematics 1579. Springer, 1994.
- [51] M. Kobylanski: Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Prob.* 28:558-602, 2000.

- [52] J.-P. Lepeltier and J.S. Martin: Existence for BSDE with superlinear- quadratic coefficient. *Stochastics and Stochastics Reports* 63:227-240, 1998.
- [53] J.-P. Lepeltier and J.S. Martin: On the existence or non-existence of solutions for certain backward stochastic equations. *Bernoulli* 8:123-137, 2002.
- [54] P. Malliavin, H. Airault, L. Kay and G. Letac: Integration and Probability. Springer, 1995.
- [55] P.A. Meyer: Martingales locales fonctionnelles additives (II). Séminaire de probabilités de Strasbourg 12:786-803, 1978.
- [56] M. Mocha and N. Westray: Quadratic semimartingale BSDEs under an exponential moments condition. Séminaire de Probabilités XLIV. Lecture Notes in Mathematics, 2046, pp 105-139. Springer, 2012.
- [57] M.-A. Morlais: Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem. *Finance Stoch.* 13:121-150, 2009.
- [58] K. Nam: Backward stochastic differential equations with superlinear drivers. PhD Thesis. Princeton University, 2014.
- [59] D. Nualart: The Malliavin Calculus and Related Topics. 2nd edition, Springer, 2006.
- [60] D. Nualart and E. Pardoux: Stochastic calculus with anticipating integrands. *Probab. Theory Relat. Fields* 78:535-581, 1988.
- [61] D. Nualart and J. Vives: Smoothness of local time and related Wiener functionals. *Potent. Anal.* 1:257-263, 1992.
- [62] E. Pardoux and S. Peng: Adapted solutions of backward stochastic equations. *System and Control Lett.* 14:55-61, 1990.
- [63] S. Peng: Backward SDE and related g -expectation. In *Backward Stochastic Differential Equations*, Pitman Research Notes in Math. No. 364, ed. by N. El Karoui and L. Mazliak, 141-159, 1997.
- [64] S. Peng: Nonlinear expectations and nonlinear Markov chains. *Chin. Ann. Math.* 26B(2):159-184, 2005.
- [65] P.E. Protter: Stochastic Integration and Differential Equations. 2nd edition, Springer, 2004.
- [66] H. Triebel: Interpolation Theory, Function spaces, Differential Operators. North-Holland, 1978.
- [67] Y. Qiu: On the UMD-constants for a class of iterated $L_p(L_q)$ spaces. *J. Funct. Anal.* 263:2409-2429, 2012.
- [68] G. Dos Reis: On some properties of solutions of quadratic growth BSDE and applications in finance and insurance. PhD Thesis, Humboldt University Berlin, 2010.
- [69] D. Revuz and M. Yor: Continuous martingales and Brownian motion. Springer, 1999.
- [70] W. Schachermayer: A characterisation of the closure of H^∞ in BMO. Séminaire de Probabilités, XXX. Lecture Notes in Mathematics, 1626, pp. 344-356. Springer, 1996.
- [71] D.W. Stroock: Applications of Fefferman-Stein type interpolation to probability theory and analysis. *Comm. Pure Appl. Math.* 26:477-495, 1973 Springer, 2011.
- [72] D.W. Stroock: Essentials of Integration Theory for Analysis. Springer, 2011.
- [73] S. Watanabe: Fractional order Sobolev spaces on Wiener space. *Probab. Theory Relat. Fields* 95:175-198, 1993.
- [74] P. Xie and X.C. Zhang: Fractional smoothness of some stochastic integrals. *Acta Mathematica Sinica* 23:1053-1058, 2007.
- [75] J. Ylinen: Decoupling on the Wiener space and variational estimates for BSDEs. PhD Thesis. Report 148 of the Department of Mathematics and Statistics, University of Jyväskylä, 2015.
- [76] J. Ylinen: Tales and tails of BSDEs. ArXiv:1501.01183.
- [77] J. Zhang: A numerical scheme for BSDEs. *Ann. Appl. Prob.* 14:459-488, 2004.
- [78] J. Zhang: Representation of solutions to BSDEs associated with a degenerate FSDE. *Ann. Appl. Prob.* 15:1798-1831, 2005.